Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 17 (2015), Issue 10 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Viewpoint

Original Paper

Health Coaching Reduces HbA1c in Type 2 Diabetic Patients From a Lower-Socioeconomic Status Community: A Randomized Controlled Trial

Noah Wayne[1], MSc, RKin, CSEP-CEP; Daniel F Perez[1], MSc; David M Kaplan[2], MD, MSc, CCFP; Paul Ritvo[1], PhD, CPsych

[1]School of Kinesiology & Health Science, Faculty of Health, York University, Toronto, ON, Canada

[2]North York General Hospital, Department of Family & Community Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Paul Ritvo, PhD, CPsych
School of Kinesiology & Health Science
Faculty of Health
York University
357 Bethune College
4700 Keele Street
Toronto, ON, M3J1P3
Canada
Phone: 1 416 736 2100 ext 22396
Fax: 1 416 736 5774
Email: pritvo@yorku.ca

Abstract

Background: Adoptions of health behaviors are crucial for maintaining good health after type 2 diabetes mellitus (T2DM) diagnoses. However, adherence to glucoregulating behaviors like regular exercise and balanced diet can be challenging, especially for people living in lower-socioeconomic status (SES) communities. Providing cost-effective interventions that improve self-management is important for improving quality of life and the sustainability of health care systems.

Objective: To evaluate a health coach intervention with and without the use of mobile phones to support health behavior change in patients with type 2 diabetes.

Methods: In this noninferiority, pragmatic randomized controlled trial (RCT), patients from two primary care health centers in Toronto, Canada, with type 2 diabetes and a glycated hemoglobin/hemoglobin A1c (HbA1c) level of ≥7.3% (56.3 mmol/mol) were randomized to receive 6 months of health coaching with or without mobile phone monitoring support. We hypothesized that both approaches would result in significant HbA1c reductions, although health coaching with mobile phone monitoring would result in significantly larger effects. Participants were evaluated at baseline, 3 months, and 6 months. The primary outcome was the change in HbA1c from baseline to 6 months (difference between and within groups). Other outcomes included weight, waist circumference, body mass index (BMI), satisfaction with life, depression and anxiety (Hospital Anxiety and Depression Scale [HADS]), positive and negative affect (Positive and Negative Affect Schedule [PANAS]), and quality of life (Short Form Health Survey-12 [SF-12]).

Results: A total of 138 patients were randomized and 7 were excluded for a substudy; of the remaining 131, 67 were allocated to the intervention group and 64 to the control group. Primary outcome data were available for 97 participants (74.0%). While both groups reduced their HbA1c levels, there were no significant between-group differences in change of HbA1c at 6 months using intention-to-treat (last observation carried forward [LOCF]) (*P*=.48) or per-protocol (*P*=.83) principles. However, the intervention group did achieve an accelerated HbA1c reduction, leading to a significant between-group difference at 3 months (*P*=.03). This difference was reduced at the 6-month follow-up as the control group continued to improve, achieving a reduction of 0.81% (8.9 mmol/mol) (*P*=.001) compared with a reduction of 0.84% (9.2 mmol/mol)(*P*=.001) in the intervention group. Intervention group participants also had significant decreases in weight (*P*=.006) and waist circumference (*P*=.01) while controls did not. Both groups reported improvements in mood, satisfaction with life, and quality of life.

Conclusions: Health coaching with and without access to mobile technology appeared to improve glucoregulation and mental health in a lower-SES, T2DM population. The accelerated improvement in the mobile phone group suggests the connectivity

provided may more quickly improve adoption and adherence to health behaviors within a clinical diabetes management program. Overall, health coaching in primary care appears to lead to significant benefits for patients from lower-SES communities with poorly controlled type 2 diabetes.

Trial Registration: ClinicalTrials.gov NCT02036892; http://clinicaltrials.gov/ct2/show/NCT02036892 (Archived by WebCite at http://www.webcitation.org/6b3cJYJOD)

(J Med Internet Res 2015;17(10):e224) doi:10.2196/jmir.4871

KEYWORDS

diabetes mellitus, type 2; health coaching; mHealth; telehealth; randomized controlled trial; RCT

Introduction

Overview

The type 2 diabetes mellitus (T2DM) epidemic is an increasing economic and personal health burden that could be cost-effectively addressed with health coach (HC) interventions, assisted by mobile phone technologies [1]. HC interventions target health behavior changes aligned with self-determined goals leading to improved physical and mental health outcomes [2]. Chronic medical conditions are targeted when health behaviors adopted by patients can significantly reduce risks of worsened disease and disease complications [3].

Amid promising reports of computer and mobile phone-assisted health interventions [4], a dearth of studies focus on which types of personal interactions combine most effectively with current technologies. Prior to this randomized controlled trial (RCT), we codeveloped, with NexJ Systems Inc, mobile phone software for logging health data (eg, blood glucose, blood pressure, mood, energy, and pain) and related activities (eg, exercise, diet, and stress) using secure, cloud-based storage. The software permits innovative comonitoring of client behaviors (eg, photographing meals) and transmission of reminder messages encouraging activation and adherence. As the HC reviews participant activities in real-time experience, these immediately responsive communications can prevent relapse and/or assist relapse recovery, as demonstrated in a pilot study [3].

Internet-based interventions have demonstrated significant improvements in glucoregulation in T2DM patients, as exemplified in a cluster RCT undertaken by Quinn et al (2011) where 4 different intensity levels of Internet-based support were compared; significant between-group differences in reduced glycated hemoglobin/hemoglobin A1c (HbA1c) were found when the most intense intervention (P<.001) was compared to usual care. This intervention consisted mainly of automated messages prompted by patient entries (eg, self-assessed blood glucose) and the patients studied were all health insured, after exclusion of the noninsured population that is often associated with lower socioeconomic status (SES), higher T2DM prevalence, and poorer glucose control [5,6]. In contrast, our intervention included a high proportion of lower-SES patients as all Ontario residents are able to access essential health services free of charge via the Ontario Health Insurance Program (OHIP). Our trial focused on supporting participants in surmounting the additional challenges confronted by lower-SES community residents, such as poor neighborhood walkability [7] and elevated consumption of energy-dense/nutrient-poor foods [8]. Failure to surmount these challenges often leads to an increased longitudinal use of health care resources due to more reactive use combined with poorer health status [9]. A further contrast was that our study was based on assessing HC interactions, with and without mobile phone-based support.

Another more recent trial compared a mobile phone-based, self-management system with and without telephone-based health coaching in improving HbA1c levels, with a usual care control group. Both intervention groups accessed a mobile phone-based self-management system that enabled users to track blood glucose, diet, physical activity, and personal goals. The most intensive intervention group received health coaching delivered by a diabetes specialist nurse for the first 4 months of the 12-month trial, with a total of five 20-minute phone contacts. Results indicated no significant between-group or within-group HbA1c differences [10]. The intensity of this HC intervention—five 20-minute phone contacts—was considerably lower than the levels applied in this study.

The importance of lowering HbA1c and improving glucoregulation in T2DM patients cannot be overemphasized as HbA1c is a robust indicator of complication risks and a widely accepted tool for T2DM diagnosis [11]. Without proper management, patients with T2DM are at increased risk for debilitating complications, particularly stroke [12], neuropathy leading to amputation and blindness [13], and death [14]. HbA1c reductions have been associated with carbohydrate control [15], vigorous exercise [16], and medication adherence [17].

While the economic pressures of funding interventions motivate technological developments that can, in part or whole, replace personal counseling interventions, studies that compare different HC intensities combined with different technologies are necessary to determine optimal proportions. The usefulness of such studies is exemplified by Nundy et al (2013) who assessed a mobile phone-based, automated text messaging and counselling intervention with type 2 diabetes patients. In a quasi-experimental, two-group, pre-/post-design, intervention participants appeared to be 8.8% less costly *during* the 6-month intervention, than during the 6 months preceding intervention engagement. These participants also reduced their HbA1c by 0.7% leading to other potential longitudinal cost savings not yet evaluated [18]. Because all were participants in the University of Chicago employee health plan, relevant health care costs were accessed and compared. Once again, our study differs in that our sample included unemployed individuals who would have been ineligible for the health plan which this previous study relied on.

The purpose of increasing the frequency and intensity of any health behavior in T2DM patients is improved glucoregulation, which directly and/or indirectly influences health-related quality of life (HRQOL). It is important to assess HRQOL outcomes independently through secondary RCT analyses, as improvements in physical health that are not associated with positive changes in quality of life are not likely to be sustainable. We used the comparison analyses of this study and, additionally, the qualitative analyses of semistructured interviews found in a companion study [19].

Objective

Based on data from a previous pilot trial, this noninferiority pragmatic RCT tested the effectiveness of a mobile phone-based health coaching protocol, versus one without mobile phone support, in reducing the HbA1c of patients with T2DM from a lower-SES community.

Methods

Overview

This pragmatic RCT proceeded with a 1:1 allocation ratio. Participants were recruited from 2 primary health clinics in Toronto, Canada, between March 2012 and October 2013. The populations served were from a lower-income neighborhood (90% of participants) and a midlevel-SES community (10% of participants). Patients were eligible for participation if diagnosed with T2DM, if they had an HbA1c ≥ 7.3% (56.3 mmol/mol) measured within 1 month of consent, and if they were under 70 years of age. Following pragmatic trial guidelines, there were no additional exclusion criteria (eg, no exclusion of individuals with psychiatric diagnoses). All study protocols were approved by the Research Ethics Boards at York University, North York Family Health Team, and North York General Hospital. This RCT was registered with ClinicalTrials.gov (NCT02036892) and reported following CONSORT-EHEALTH statement guidelines [20].

Recruitment was undertaken through phone contacts with eligible individuals identified via clinic electronic medical records. Additional recruitment assistance was obtained from associated diabetes education programs, primary care physicians, and locally practicing endocrinologists.

When participants agreed to an initial meeting to discuss the study, their HbA1c findings were verified, the study protocol was explained, and informed consent was obtained. Eligible patients then completed demographic and psychometric questionnaires and were randomized. Table 1 shows the baseline characteristics of the participants.

Table 1. Baseline characteristics (as per study protocol).

Baseline characteristics	Whole sample (n=97), mean (SD) or n (%)[a]	Intervention group (n=48), mean (SD) or n (%)[a]	Control croup (n=49), mean (SD) or n (%)[a]
Age in years, mean (SD)	53.2 (11.3)	53.1 (10.9)	53.3 (11.9)
Location, n (%)			
Site #1: BCCHC[b]	90 (93)	46 (96)	44 (90)
Site #2: NYFHT[c]	7 (7)	2 (4)	5 (10)
Gender, n (%)			
Male	27 (28)	17 (35)	10 (20)
Female	70 (72)	31 (65)	39 (80)
Ethnicity, n (%)			
First Nations	1 (1)	0 (0)	1 (2)
Black: African	5 (5)	3 (6)	2 (4)
Black: Caribbean	39 (40)	19 (40)	20 (41)
Caucasian	26 (27)	12 (25)	14 (29)
Hispanic	9 (9)	4 (8)	5 (10)
South Asian	4 (4)	3 (6)	1 (2)
South East Asian	4 (4)	2 (4)	2 (4)
West Indian	6 (6)	3 (6)	3 (6)
Other	3 (3)	2 (4)	1 (2)
Highest education level achieved, n (%)			
Less than high school	22 (23)	10 (21)	12 (24)
High school diploma	35 (36)	17 (35)	18 (37)
College or vocational training	25 (26)	11 (23)	14 (29)
University degree	12 (12)	8 (17)	4 (8)
Not disclosed	3 (3)	2 (4)	1 (2)
Employment, n (%)			
Unemployed	35 (36)	16 (33)	19 (39)
Student	4 (4)	3 (6)	1 (2)
Part time	6 (6)	1 (2)	5 (10)
Full time	25 (26)	13 (27)	12 (25)
Retired	11 (11)	6 (13)	5 (10)
Self-employed	9 (9)	6 (13)	3 (6)
Work in home (eg, take care of children)	4 (4)	2 (4)	2 (4)
Not disclosed	3 (3)	1 (2)	2 (4)
Income in Can $, n (%)			
$0-$9999	21 (22)	9 (19)	12 (25)
$10,000-$25,000	23 (24)	10 (21)	13 (27)
$25,000-$50,000	20 (21)	12 (25)	8 (16)
$50,000-$75,000	9 (9)	3 (6)	6 (12)
$75,000 and higher	5 (5)	4 (8)	1 (2)
Not disclosed	19 (20)	10 (21)	2 (4)

Baseline characteristics	Whole sample (n=97), mean (SD) or n (%)[a]	Intervention group (n=48), mean (SD) or n (%)[a]	Control croup (n=49), mean (SD) or n (%)[a]
Car access, n (%)			
Owns a car	35 (36)	19 (40)	16 (33)
Has access to car	12 (12)	9 (19)	3 (6)
No access to car	48 (50)	19 (40)	29 (59)
Not disclosed	2 (2)	1 (2)	1 (2)

[a]Percentages may not add up to 100% due to rounding.

[b]Black Creek Community Health Centre (BCCHC).

[c]North York Family Health Team (NYFHT).

Intervention

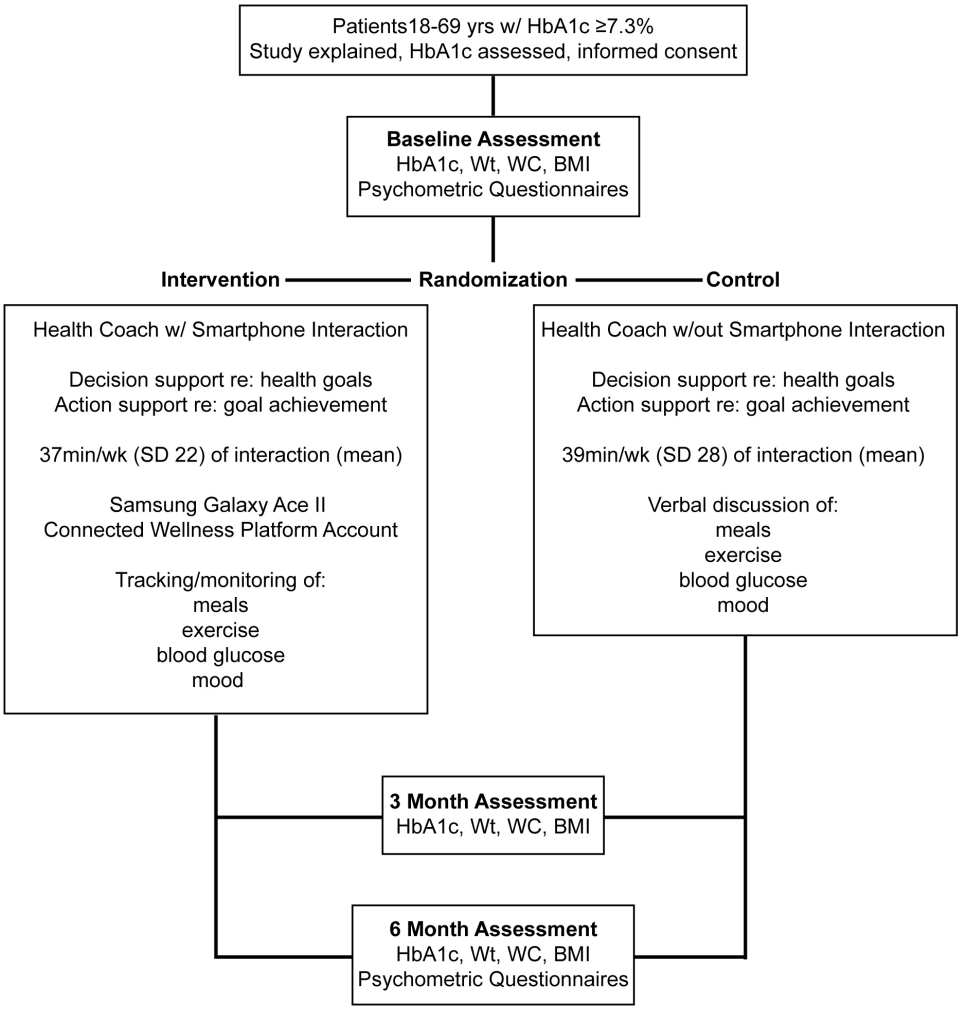
The HC intervention extended for 6 months from the date of consent (Figure 1) following a behavior-change curriculum designed by 2 study authors (PR and NW) at York University that incorporated feedback from the prior pilot study [3]. In the intervention, a health coach was defined as a behavior-change counselling specialist with expertise in chronic disease management and evidence-based theory adapted for disease state, SES, and ethnocultural backgrounds. With HC assistance, clients determined health-related goals and monitored daily progress. The HC comonitored the client's mobile phone input and directed immediate attention (on a 24-hour/day and 7-day/week basis) to episodes of desirable progress, relapse, and resistance. The HC protocol has been manualized, emphasizing those situations observed to frequently arise when behavior change is addressed in T2DM-affected individuals.

Eligible participants were randomized to the respective study groups (with and without mobile phone support), with HCs in both groups guiding participants in planning and reaching health targets aimed at reducing HbA1c. Efforts focused primarily on increasing exercise (frequency, duration, intensity) and modifying diet to reduce carbohydrate intake. Additional goals emphasized stress management, medication adherence, and effective communication with primary care physicians and, generally, within the health system.

Six HCs intervened with experimental and control group participants. These individuals held bachelor's degrees in kinesiology and health science and/or were graduate students in the School of Kinesiology and Health Science at York University. Five HCs were certified exercise physiologists and one was a certified personal trainer, all certified by the Canadian Society for Exercise Physiology (CSEP). All attended weekly seminars prior to and throughout the trial where they received training in the HC curriculum by the lead investigator (PR). HCs also participated in weekly team meetings led by the study coordinator (NW) where they discussed applications of behavior theory in specific strategies for each participant.

The Black Creek Community Health Centre (BCCHC) concurrently provided the Exercise Education Program (EEP) to all community members (free of charge) that featured exercise prescription, monitoring, and adherence support. Participants were monitored on both an individual and group basis by trainers during exercise sessions and patients with T2DM were provided with special blood glucose testing before and after each exercise session. The program included group exercise classes, resistance training with weights and bands, and cardiovascular exercise using a treadmill and stationary bicycles. Both intervention and control group participants had EEP access for the trial duration.

Figure 1. Experimental design and timing of data collection.



Intervention Group

The intervention group was provided with a Samsung Galaxy Ace II mobile phone running Google Android Ice Cream Sandwich (Android 4.0.2) for the study intervention period, with a data-only carrier plan. They were also provided a user account with the Connected Wellness Platform (CWP) provided by NexJ Systems, Inc [21], which supported participants in health-related goal setting and progress monitoring. Participants could track key metrics, notably blood glucose levels (Figure 2), exercise frequency/duration/intensity (Figure 3), food intake (via photo journaling) (Figure 4), and mood (Figure 5). They could communicate with their health coach at any time in the 24-hour cycle via secure messaging, scheduled phone contact, and/or during in-person meetings. The mean total contact (for all these activities) was 38 minutes/week (SD 25). All health data entered by participants into the CWP were immediately visible to health coaches through a secure, Web-accessible portal. Although participants were encouraged to use the system daily, individual usage patterns varied. Participant data and software-enabled communication required two-way, certificate-based authentication and passwords that were stored in encrypted columns. The CWP exceeds Canadian privacy standards for software carrying health information. Based on patient goals, HCs used the 24-hour/day logging function to guide healthy lifestyle choices, while providing support when clients diverged from intended health goals and routines.

Figure 2. Screenshot of blood glucose tracker on the Connected Wellness Platform from NexJ Systems, Inc.

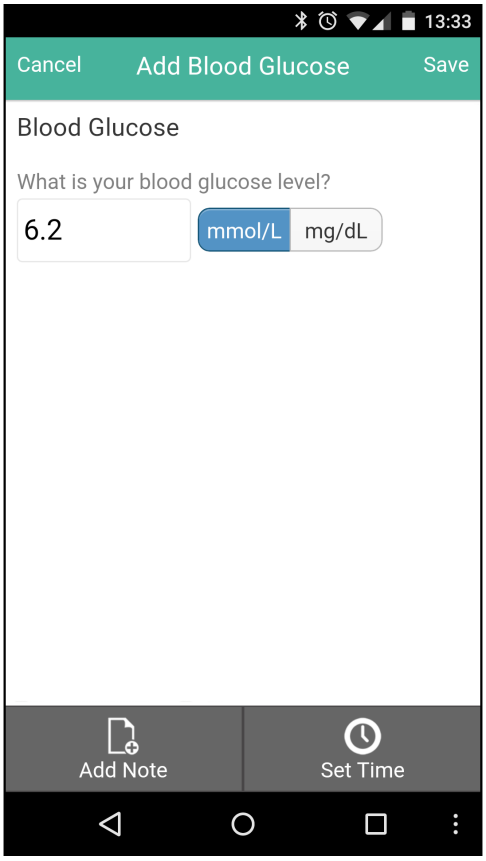


Figure 3. Screenshot of exercise tracker on the Connected Wellness Platform from NexJ Systems, Inc.

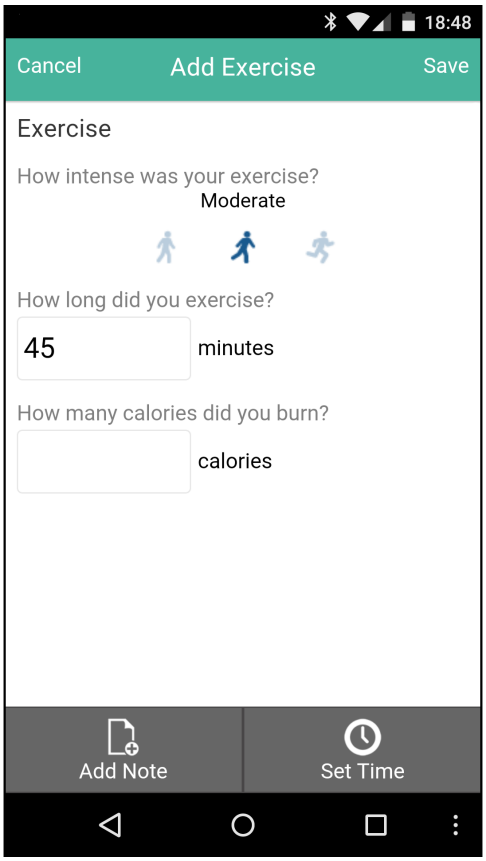


Figure 4. Screenshot of food tracker on the Connected Wellness Platform from NexJ Systems, Inc.

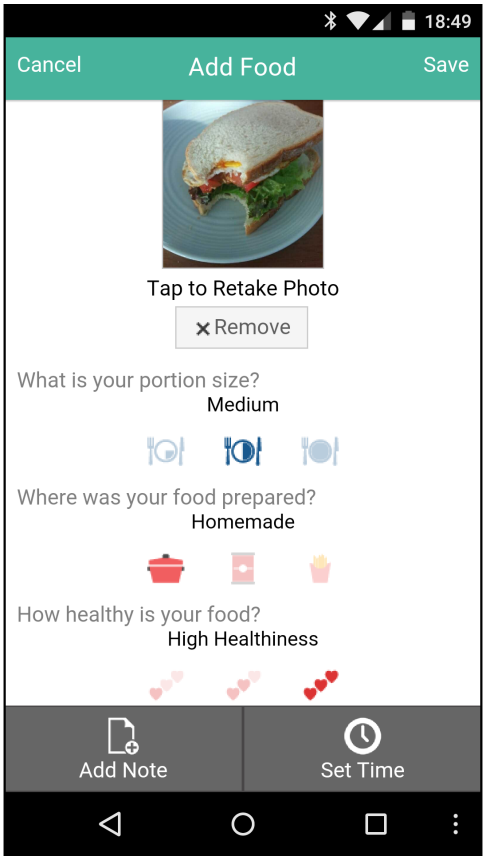


Figure 5. Screenshot of mood tracker on the Connected Wellness Platform from NexJ Systems, Inc.

Control Group

Control group participants received HC support in selecting and progressing toward goals without access to a (study-provided) mobile phone or the CWP software. Control group participants accessed the EEP, as did the intervention group participants for the study duration, in addition to in-person meetings and health coach phone contacts.

Primary Outcome

The primary outcome was the difference between intervention and control group means of HbA1c levels from baseline to 6 months. Intention-to-treat and per-protocol analyses were both undertaken and are presented below. HbA1c levels were assessed by physician requisition or, when unobtainable, by a point-of-care HbA1c analyzer (Siemens DCA Vantage 3000) which has met performance criteria in efficacy trials [22] and has been employed in comparable research trials [10,23]. To ensure consistency, the type of HbA1c collection at baseline was the same at follow-up sessions. While the 3-month assessment allowed an evaluation of trends, the 6-month assessment was used as the primary outcome. Measures of blood work were accepted within 4 weeks of the 3- and 6-month measurement intervals providing flexibility for participant schedules and physician requisitions.

Secondary Outcomes

Differences between HbA1c mean levels within groups were also analyzed. Additional outcomes included anthropometric measurements for weight (kg), body mass index (BMI) (kg/m^2), and waist circumference (cm) collected at baseline and 6-month time points. Changes in psychometric assessments at baseline and 6 months were analyzed using the Satisfaction with Life Scale [24], the Hospital Anxiety and Depression Scale [25], the Positive and Negative Affect Schedule [26], and the Short Form Health Survey-12 (SF-12) [27]. All measures were obtained on-site by research staff.

Sample Size

An a priori power calculation indicated that 48 participants were needed per group to detect an estimated difference of HbA1c of 0.65%, assuming a significance level of 5% (two-tailed), a standard deviation of 1.4, and a statistical power of 80%. We overenrolled to allow for attrition, setting our final recruitment target at 65 participants per group.

Randomization

A random number sequence was generated using a random number-generating program without constraints [28]. After the sequence was generated by the research coordinator, a research assistant with no connection to the trial sealed the sequence in individual, opaque envelopes and numbered each based on sequence generation. Once a candidate participant consented and their HbA1c was verified as meeting the inclusion criteria, the next envelope was opened (in sequence) to ascertain group allocation, and the health coaching intervention commenced. Patient and coach blinding was impossible as participants readily identified receipt of a mobile phone with experimental group participation and the absence of receipt with control group participation.

Statistical Analysis

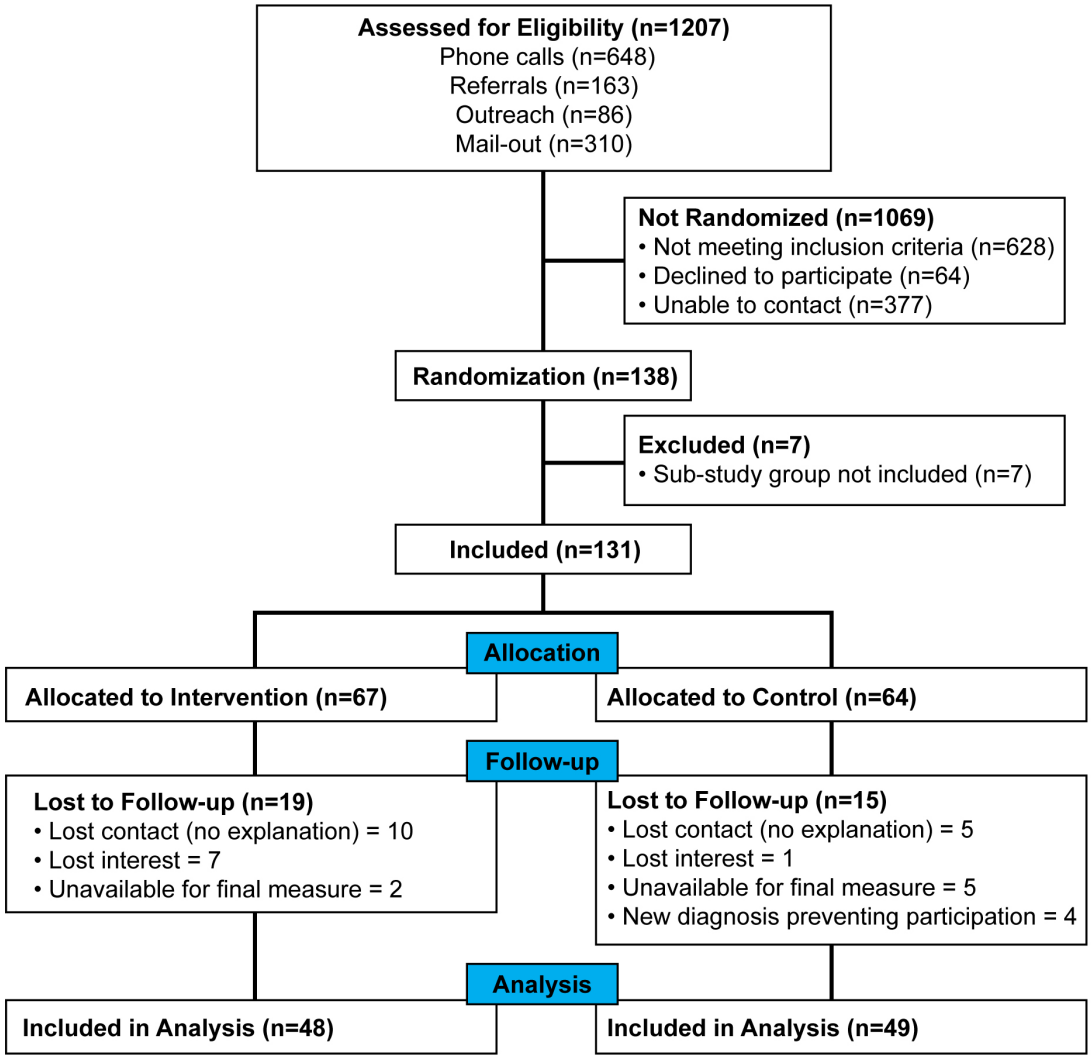
Data were double entered by 2 independent research assistants to ensure accuracy. Baseline characteristics between intervention and control groups were compared for differences using independent samples *t* tests for continuous variables and chi-square for dichotomous variables. Primary outcome comparison was conducted with an independent samples *t* test using per-protocol and intention-to-treat analyses (last observation carried forward [LOCF]). Secondary outcome comparisons were conducted solely using per-protocol comparisons with a factorial repeated-measures analysis of variance (ANOVA). Data were analyzed using SPSS 21.0 (IBM Corp, Armonk, NY, USA).

Results

Overview

Between March 2012 and October 2013, 138 participants were recruited; 67 were randomized to the experimental arm and 64 to the control arm (7 were excluded for substudy analysis) as seen in the CONSORT diagram (Figure 6). A majority of participants (57/97, 59%) had not completed postsecondary education and 35 out of 97 (36%) were unemployed, while a total of 64 out of 97 (66%) reported household incomes of Can $50,000 or less. A majority of participants were recruited from Site Number 1 (90/97, 93%) and were female (70/97, 72%). Of the 131 participants included in the study, there were 34 dropouts (26%), with 19 out of 67 (28%) from the intervention group and 15 out of 64 (23%) from the control group. Independent samples *t* tests indicated no statistically significant differences between dropouts and trial completers for HbA1c or for demographic variables. Final per-protocol analysis included 97 participants out of 131 (74%), with 48 in the intervention group and 49 in the control group. Of the measures collected, differences at baseline between groups were only detected for the SF-12 Mental Health Composite Scores. Of the 48 participants allocated to the mobile phone group, mobile phone use data indicated that 39 out of 48 participants (81%) used the CWP with consistency to communicate with their health coach and track various health measures (eg, blood glucose, food, and/or exercise).

Figure 6. Flowchart of enrollment.



Hemoglobin A1c

Independent samples *t* tests indicated no significant between-group differences in HbA1c from baseline to 6 months when analyzed with intention-to-treat (P=.48) and per-protocol (P=.83) principles (Table 2).

Table 2. Independent samples *t* test measuring differences in HbA1c levels from baseline to 6 months.

Type of analysis	n	Intervention group, mean (SD)	Control group, mean (SD)	Difference	*P* (two-tailed)
HbA1c: per protocol	97	-0.815 (1.050)	-0.759 (1.390)	0.055	.83
HbA1c: intention to treat	131	-0.642 (1.040)	-0.974 (1.400)	0.152	.48

Results from a repeated-measures ANOVA indicated trends for between-group HbA1c differences in a per-protocol analysis—$F_{1,89}$=3.022, P=.09 (Table 3).

Table 3. Between-group analysis of variance measuring differences in HbA1c levels.

Type of analysis	n	Type II sum of squares	df	Mean square	$F_{1,89}$	*P*	Partial eta squared
HbA1c: per protocol	97	3.004	1	3.004	3.002	.09	.034
HbA1c: intention to treat	131	1.463	1	1.463	1.142	.29	.009

These differences reflected significant HbA1c within-group reductions from baseline to 6 months in the intervention group—0.84% (9.2 mmol/mol), 95% CI 0.46-1.17; *P*=.001—and in the control group—0.81% (8.9 mmol/mol), 95% CI 0.41-1.11; *P*=.001—(Table 4), and a significantly greater reduction for the intervention group versus the control group at the 3-month follow-up (*P*=.03; Table 5).

Table 4. Change in HbA1c levels by group.

Measurement time point	Intervention group			Control group		
	n	Mean % (SD or 95% CI)	Total value (mmol/mol)	n	Mean % (SD or 95% CI)	Total value (mmol/mol)
HbA1c included in *t* test (n=97)						
Baseline, mean (SD)	48	8.69 (1.32)	71.5	49	8.89 (1.30)	73.7
6 months, mean (SD)	48	7.88 (1.17)	62.6	49	8.13 (1.27)	65.4
Change from baseline to 6 months, mean (95% CI)	48	0.82 (0.46-1.17)[a]	8.9	49	0.76 (0.41-1.11)[a]	8.3
HbA1c included in ANOVA [b] (n=89)						
Baseline, mean (SD)	45	8.60 (1.19)	70.5	44	8.88 (1.32)	73.6
3 months, mean (SD)	45	7.74 (1.06)	61.1	44	8.26 (1.16)	66.8
6 months, mean (SD)	45	7.76 (1.00)	61.3	44	8.07 (1.29)	64.7
Change from baseline to 3 months, mean (95% CI)	45	0.86 (0.47-1.26)[a]	9.4	44	0.62 (0.23-1.03)[a]	6.8
Change from baseline to 6 months, mean (95% CI)	45	0.84 (0.38-1.26)[a]	9.2	44	0.81 (0.34-1.28)[a]	8.9

[a]Significant at the *P*=.001 level.

[b]Analysis of variance (ANOVA).

A data discrepancy was detected during the repeated-measures ANOVA as 3 participants in the intervention group and 5 in the control group were not assessed at 3 months but were evaluated at 6 months. They had either refused the 3-month testing or their family physicians failed to provide their test results. Subsequent *t* tests indicated a lesser reduction in HbA1c (baseline to 6 months) for the controls lacking the 3-month data versus completers (*P*=.03). There were no differences in HbA1c (baseline to 6 months) for intervention participants lacking 3-month data versus those with complete data. Furthermore, no significant differences were found in baseline HbA1c levels for either intervention or controls participants with or without a 3-month HbA1c measure.

Table 5. HbA1c values for participants with and without 3-month measurements.

Measurement	Intervention group			Control group		
	3-month measure absent (n=3)	3-month measure present (n=45)	*P*	3-month measure absent (n=5)	3-month measure present (n=44)	*P*
Baseline HbA1c, mean % (SD)	9.97 (2.64)	8.60 (1.19)	.47	8.92 (1.19)	8.88 (1.33)	.95
Total HbA1c value (mmol/mol)	85.5	70.5		74	73.6	
Change in HbA1c (6 month-baseline), mean % (SD)	-0.40 (0.46)	-0.84 (1.08)	.47	-0.28 (0.19)	-0.81 (1.45)	.03

Table 6 shows that the HbA1c trend differences indicated with the repeated-measures ANOVA—$F_{1,89}$=3.022, *P*=.09—were due to the greater reduction of HbA1c at 3 months in the intervention versus control group. This between-group difference disappeared at 6 months with gains in the control group, and no further gains in the intervention group.

Table 6. Time-point comparison of HbA1c levels for intervention versus control groups.

Time point	Between-group difference of % HbA1c (95% CI)	*P*
Baseline	0.280 (-0.250 to 0.810)	.30
3 months	0.515 (0.500 to 0.990)	.03
6 months	0.308 (-0.180 to 0.800)	.21

Secondary Outcomes: Body Composition

We detected significant reductions in body weight (1.22 kg, 95% CI 0.35-2.08; *P*=.006) and waist circumference (2.23 cm, 95% CI 0.53-3.93; *P*=.01) in the intervention group, while the control group had no change. There were no significant changes in BMI in either group (Table 7).

Secondary Outcomes: Psychometric Questionnaires

A significant number of trial completers chose not to complete psychometric questionnaires at follow-up, resulting in their baseline outcomes being omitted from additional analyses (Table 7). Comparison of the baseline psychometric outcomes of completers and noncompleters indicated no significant differences.

Within-group, pre/post improvements in life satisfaction were detected in the intervention (+3.72, 95% CI 1.50-5.94; *P*=.001) and control groups (+3.77, 95% CI 1.30-6.24; *P*=.003) (Satisfaction with Life Scale). Similar improvements for both intervention and control groups were detected in the Hospital Anxiety and Depression Scale (HADS) depression subscale (-1.81, 95% CI -2.81 to -0.81; *P*=.001; -1.70, 95% CI -2.73 to -0.67; *P*=.002), and the Physical Composite Score of the SF-12 (+2.69, 95% CI 0.21-5.17; *P*=.03; +2.92, 95% CI 0.24-5.60; *P*=.03) (Table 7), although the control group demonstrated a significantly reduced HADS anxiety subscale score (-1.50, 95% CI -2.73 to -0.27; *P*=.02), while the intervention group did not (-1.12, 95% CI -2.29 to 0.05; *P*=.06) (Table 7). Significant between-group differences were found at the 6-month follow-up for negative affect (negative affect subscale of the Positive and Negative Affect Schedule [PANAS]) (+5.27, 95% CI 1.51-9.04; *P*=.007) favoring the intervention group (Table 7).

Table 7. Baseline, follow-up, and change values for all secondary outcomes.

Variable by group	n	Baseline, mean (SD or 95% CI)	6-month follow-up, mean (SD or 95% CI)	Change, mean (95% CI)	*P*
Weight (kg)					
Intervention	41	93.66 (20.23)	92.44 (20.24)	-1.22 (0.35-2.08)[a]	.006
Control	39	98.76 (24.02)	99.21 (24.77)	+0.45 (-1.33 to 0.44)	.32
Difference between groups		5.10 (-4.78 to 14.98)	6.76 (-3.29 to 16.81)		
P		.31	.18		
Waist circumference (cm)					
Intervention	40	112.11 (14.50)	109.88 (14.82)	-2.23 (0.53-3.93)[a]	.01
Control	37	113.88 (17.04)	114.00 (18.12)	+0.122 (-1.89 to 1.64)	.89
Difference between groups		1.78 (-5.39 to 8.94)	4.13 (-3.36 to 11.62)		
P		.62	.28		
Body mass index (kg/m ^{2})					
Intervention	39	33.74 (6.70)	33.53 (6.80)	-0.21 (-0.24 to 0.66)	.35
Control	36	37.00 (7.92)	37.21 (8.22)	-0.21 (-0.68 to 0.25)	.37
Difference between groups		3.26 (-0.11 to 6.63)	3.69 (0.22-7.15)[a]		
P		.06	.04		
Satisfaction with Life					
Intervention	32	20.50 (7.71)	24.22 (6.33)	+3.72 (1.50-5.94)[b]	.001
Control	26	18.04 (7.01)	21.81 (7.15)	+3.77 (1.30-6.24)[b]	.003
Difference between groups		-2.46 (-1.46 to 6.38)	-2.41 (-1.14 to 5.96)		
P		.21	.18		
HADS [c] : anxiety subscale					
Intervention	33	7.39 (4.53)	6.27 (4.18)	-1.12 (-2.29 to 0.05)	.06
Control	30	9.50 (4.49)	8.00 (5.06)	-1.50 (-2.73 to -0.27)[a]	.02
Difference between groups		2.11 (-0.17 to 4.39)	1.73 (-0.60 to 4.06)		
P		.07	.14		
HADS: depression subscale					
Intervention	32	6.25 (3.99)	4.44 (3.32)	-1.81 (-2.81 to -0.82)[b]	.001
Control	30	7.77 (4.06)	6.07 (4.38)	-1.70 (-2.73 to -0.67)[b]	.002
Difference between groups		1.52 (-0.53 to 3.56)	1.63 (-0.34 to 3.60)		
P		.14	.10		
PANAS [d] : positive affect subscale					
Intervention	30	34.43 (8.46)	36.03 (7.65)	+1.60 (-1.00 to 4.20)	.22
Control	27	31.22 (10.29)	31.67 (9.71)	+0.44 (-2.30 to 3.18)	.75
Difference between groups		-3.21 (-1.77 to 8.19)	-4.37 (-0.25 to 8.98)		
P		.20	.06		
PANAS: negative affect subscale					
Intervention	31	16.58 (7.85)	14.55 (5.03)	-2.03 (-4.87 to 0.80)	.16
Control	28	20.39 (9.57)	19.82 (9.04)	-0.57 (-3.55 to 2.41)	.70
Difference between groups		3.81 (-0.73 to 8.36)	5.27 (1.51-9.04)[a]		

Variable by group	n	Baseline, mean (SD or 95% CI)	6-month follow-up, mean (SD or 95% CI)	Change, mean (95% CI)	*P*
P		.10	.007		
SF-12 [e] : Physical Composite Score					
Intervention	34	42.89 (8.69)	45.57 (7.54)	+2.69 (0.21-5.17)[a]	.03
Control	29	41.63 (10.08)	44.55 (10.89)	+2.92 (0.24-5.60)[a]	.03
Difference between groups		1.25 (-3.48 to 5.98)	1.02 (-3.65 to 5.68)		
P		.60	.66		
SF-12: Mental Composite Score					
Intervention	34	47.74 (11.11)	50.22 (10.29)	+2.48 (-1.10 to 6.05)	.17
Control	29	41.68 (11.82)	44.50 (10.15)	+2.82 (-1.05 to 6.69)	.15
Difference between groups		6.06 (0.28-11.85)[a]	5.72 (0.56-10.89)[a]		
P		.04	.03		

[a]The change is significant, $P<.05$.

[b]The change is significant, $P<.005$.

[c]Hospital Anxiety and Depression Scale (HADS).

[d]Positive and Negative Affect Schedule (PANAS).

[e]Short Form Health Survey-12 (SF-12).

Discussion

Principal Findings

Personalized health coaching with and without the provisions of mobile phone and related software support was assessed in a predominantly lower-SES population with poorly controlled T2DM. A total of 45% of participants reported household incomes of Can $25,000 or less, qualifying them as living at or beneath the Canada poverty line [29], while an additional 20.9% of participants reported household incomes between Can $25,000 and Can $50,000. Our findings suggest clinically significant within-group reductions in HbA1c in both groups but no significant between-group differences in HbA1c from baseline to 6 months according to per-protocol (P=.83) and intention-to-treat (LOCF) (P=.48) analyses.

There was, however, a significant between-group difference in HbA1c at the 3-month time point (0.52%, P=.03) favoring the mobile phone-assisted group, although this difference was not statistically significant at 6 months because the control group's mean HbA1c reduction improved between 3 and 6 months while the intervention group's HbA1c level remained stable (Figure 7). This result indicates that clinically significant HbA1c reductions occurred at a faster rate with HC and mobile phone support than with solely HC support. The repeated-measures ANOVA analysis of three time points was affected by missing data; however, all missing control participants had no HbA1c reductions, resulting in an increased mean difference in remaining controls necessitating a larger effect size in the experimental condition to reflect a significant difference. Observed weight and waist circumference differences also suggested comparative benefits for the mobile phone-assisted group versus controls. These included significant reductions in weight and waist circumference in the mobile phone group which appeared to be related to the food photo-journaling function of the CWP. By reviewing photographs of their meals, participants could reflect on portion size and nutritional value in discussion *with* the health coach. These photo-stimulated "teachable moments" appeared to improve dietary choices more than was evident in the health coach-only group. Those in the mobile phone group also subjectively reported value in photographing meals and recording glucose levels in response to in-depth semistructured interviews [19]. Reductions in negative affect are likely linked to intervention participants feeling fundamentally connected in their health-focused program as their mobile phone became a constant symbol of being able to access a genuinely concerned person (24 hours a day/7 days a week) whose sole purpose was to help address health concerns. This feeling of health coach connectedness was a principal theme in the qualitative analyses of participant interviews [19].

Lower-SES populations confront higher mortality risks than equivalent higher-SES populations [6]. Due to a variety of challenges to health maintenance, individuals from lower-SES communities have poorer health status *and* use health care services more reactively [9]. They are also more likely to suffer from mental health conditions [30], *but* less likely to access mental health service resources [31]. Our results indicate that psychological well-being within the overall sample improved from baseline to 6-month follow-up, specifically demonstrated in outcomes on the Satisfaction with Life and the Hospital Anxiety and Depression Scales. As both groups communicated at least a once per week with their HCs, these interactions appeared sufficient for improvements in self-reported mood. Although differences in our primary outcome (HbA1c level) were only trending toward significant between-group differences, significant differences appeared in other markers

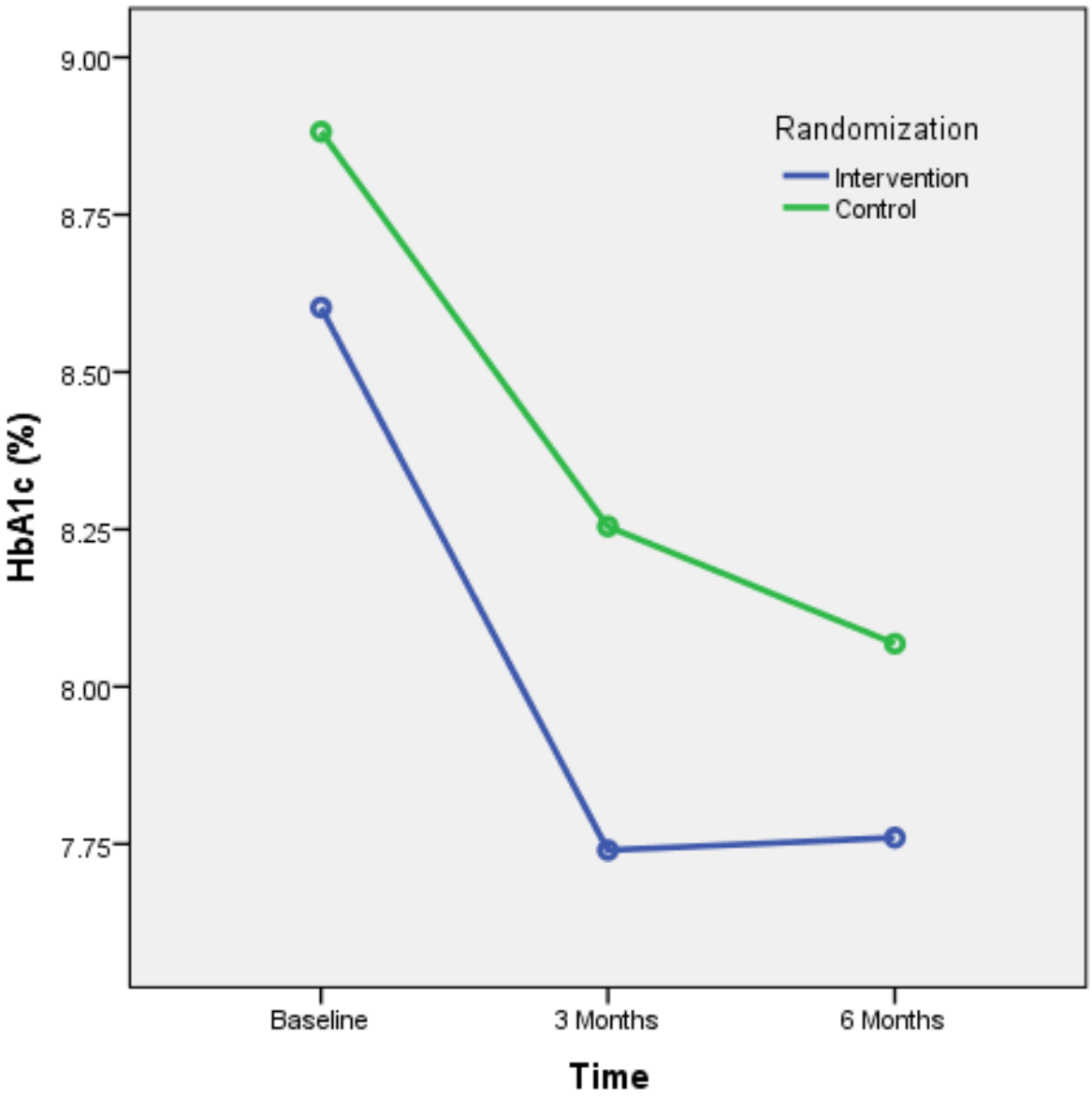
of basic health (ie, weight and waist circumference), and in the negative affect subscale of the PANAS. Once again, those who used the mobile phone subjectively reported value in photographing meals and recording blood glucose levels when responding to in-depth semistructured interviews [19].

The Connected Wellness Platform enabled self-monitoring and health coach interactions with intervention participants, providing a cloud-based platform for mobile phone-based health management. This system provided secure, two-way communication between client and health coach, with mobile phone data entry on relevant behaviors entered manually. While the restriction to manual entry was not ideal, Bluetooth functionality for glucometers and pedometers was not yet integrated into the system during the trial. Other chronic disease management systems with similar features have been tested for usability and functionality. Notably, Martinez-Millana et al [32] comprehensively tested a diabetes management system with 30 patients and assessed the speed accuracy of tracking with several Bluetooth-enabled devices (ie, glucometers and pedometers) and their performance with a variety of mobile phones and network connections. Although we did not focus on the same performance analysis criteria during this trial, the CWP went through multiple upgrades during the pilot trial [3], ensuring smoother functionality and a more refined user interface (Figures 2-5) for the RCT. Detailed user experience with the CWP was collected using semistructured interviews and is reported in a full-length article [19]. CWP-user data logs were also extracted and analyzed with data mining methods to evaluate more finely tuned associations between app use and clinical outcomes (in a submitted manuscript).

Careful titrations of health coach interventions, typically measured by the frequency and duration of patient-coach interactions, are important elements in determining the optimal HC contact for eliciting improved health at minimal cost. With too little interaction, HC interventions risk insignificant or unsustainable health improvements, while too much interaction results in overly expensive implementation. As such, studies using multiple intervention intensities are necessary to ultimately determine appropriate contact level. Although we did not specify a minimum-maximum intervention intensity during the trial (providing weekly contact was maintained), the mean interaction intensity was 38 min/week (SD 25). In both intervention and control conditions, significant improvements in HbA1c levels and psychological functioning were found. The mobile technology appeared useful in engaging participants more quickly such that significantly greater HbA1c reductions were evident at 3 months (compared to controls), which may have cost-effectiveness implications as these gains were stabilized and evident at 6 months, although additional improvements in controls ultimately erased the 3-month differences. While the gains made at 3 months were sustained at 6 months (in the intervention group), there is no evidence that gains made in either group were sustained beyond the 6-month follow-up.

A unique feature of this study was the enhanced usual care that at least partly explains gains achieved by both control and intervention participants. The BCCHC site maintained a clinical exercise program that was several yards from the primary care physician and diabetes education team offices, symbolizing the importance of exercise in health maintenance, while serving patients in need. Moreover, the program provided T2DM patients with education, exercise prescription, and monitoring, which included the assessment of blood glucose levels before and after every supervised exercise session. This supported patients in recognizing the benefits of exercise in blood glucose regulation, and helped encourage adoption of home-based exercise programs. Since the HCs in this trial were all certified exercise specialists (through the Canadian Society for Exercise Physiology), exercise prescription was undertaken safely, with no adverse events, according to the highest evidence-based standards. A total of 23 intervention patients and 22 control patients participated in the Exercise Education Program. Although we might have included a control condition that did not access the EEP, the EEP was adopted as usual care at BCCHC and denial of access would have been unethical.

Figure 7. HbA1c levels for the control and intervention groups over time.



Limitations

As with any behavioral intervention, motivations to participate introduce potential biases as those who met inclusion criteria but declined to participate represent an unstudied population. This limits the generalizability of the intervention [33]. As well, the comparison group received health coach support (without mobile monitoring) as opposed to usual care. Not only did this enable a more clear understanding of the effect of electronic monitoring of health behavior on clinical outcomes, pilot trial findings suggested a usual care control condition (ie, no health coaching) would result in an unacceptably high attrition rate in the controls. The lack of between-group differences at 6 months may be due to other, more complex factors. For example, since health coaches were randomly assigned to participants in both arms, it is possible that more effort was expended in coaching the mobile phone-assisted arm. However, since the effect size of HbA1c reduction was similar across groups, this was unlikely. Furthermore, there could have been bias in the opposite direction, with health coaches expending more effort in assisting the behavior change of control participants since these controls did not have the support of the mobile phone interactions. Also, although it would have been ideal to compare multiple glucose measures (eg, random blood sugar, fasting blood sugar), it was not possible at the participating sites. We were limited to reliable access only to HbA1c blood tests. We recognize, with other researchers, that glucose regulation is more complex than what is solely indicated in HbA1c assessment.

Conclusions

Although this trial did not indicate a significant between-group difference in improved glucoregulation, there were overall clinical and statistically significant improvements in HbA1c for participants in both health-coached groups. Given the pragmatic trial design, our findings suggest health coaching in primary care can improve the glucose management of poorly controlled

T2DM in lower-SES community residents. It is evident that using mobile phones to further connect patients to health coaches and monitor health behaviors can lead to faster reductions in HbA1c, which may have specific benefits for cost savings and quality of life. Further research comparing health-coaching interventions of different contact intensities, using wearable biomonitoring devices, and using a true waitlist/control group will help evaluate health coach intervention effectiveness, as well as long-term adherence levels and cost/benefit results.

Acknowledgments

Funding was obtained through the Public Health Agency of Canada (Project #0690490). Additional funding was obtained from York University (Connected Health and Wellness Project) through the Federal Development Agency of Southern Ontario. We would also like to thank our community partners for their support and trust which made this trial possible, most specifically Ms Michelle Westin, Ms Cheryl Prescod, Dr Ali Torshizi, and Dr Howard To.

Conflicts of Interest

None Declared.

Authors' Contributions

NW was responsible for study conception and design; data acquisition, analysis, and interpretation; and for drafting and critical revision of the manuscript. DP was responsible for data acquisition, analysis, and interpretation; and for drafting and critical revision of the manuscript. DMK was responsible for study design; data acquisition, analysis, and interpretation; and for drafting and critical revision of the manuscript. PR was responsible for study conception and design; data acquisition, analysis, and interpretation; drafting and critical revision of the manuscript; and for obtaining and administering funding.

Multimedia Appendix 1

CONSORT-eHEALTH (V 1.6.1) - submission/publication Form.

[PDF File (Adobe PDF File), 991KB - jmir_v17i10e224_app1.pdf]

References

1. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
2. Huffman M. Health coaching: a new and exciting technique to enhance patient self-management and improve outcomes. Home Healthc Nurse 2007 Apr;25(4):271-274. [doi: 10.1097/01.NHH.0000267287.84952.8f] [Medline: 17426499]
3. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]
4. Moore JO, Marshall MA, Judge DC, Moss FH, Gilroy SJ, Crocker JB, et al. Technology-supported apprenticeship in the management of hypertension: a randomized controlled trial. J Clin Outcomes Manag 2014 Mar;21(3):110-122 [FREE Full text]
5. Dinca-Panaitescu S, Dinca-Panaitescu M, Bryant T, Daiski I, Pilkington B, Raphael D. Diabetes prevalence and income: results of the Canadian Community Health Survey. Health Policy 2011 Feb;99(2):116-123. [doi: 10.1016/j.healthpol.2010.07.018] [Medline: 20724018]
6. Saydah SH, Imperatore G, Beckles GL. Socioeconomic status and mortality: contribution of health care access and psychological distress among U.S. adults with diagnosed diabetes. Diabetes Care 2013 Jan;36(1):49-55 [FREE Full text] [doi: 10.2337/dc11-1864] [Medline: 22933434]
7. Booth GL, Creatore MI, Moineddin R, Gozdyra P, Weyman JT, Matheson FI, et al. Unwalkable neighborhoods, poverty, and the risk of diabetes among recent immigrants to Canada compared with long-term residents. Diabetes Care 2013 Feb;36(2):302-308 [FREE Full text] [doi: 10.2337/dc12-0777] [Medline: 22988302]
8. Konttinen H, Sarlio-Lähteenkorva S, Silventoinen K, Männistö S, Haukkala A. Socio-economic disparities in the consumption of vegetables, fruit and energy-dense foods: the role of motive priorities. Public Health Nutr 2013 May;16(5):873-882. [doi: 10.1017/S1368980012003540] [Medline: 22857602]
9. Alter DA, Stukel T, Chong A, Henry D. Lesson from Canada's Universal Care: socially disadvantaged patients use more health services, still have poorer health. Health Affairs 2011 Feb;30(2):274-283 [FREE Full text] [doi: 10.1377/hlthaff.2009.0669] [Medline: 21289349]
10. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian

randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
11. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee. Canadian Diabetes Association 2013 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada. 2013. URL: http://guidelines.diabetes.ca/app_themes/cdacpg/resources/cpg_2013_full_en.pdf [accessed 2015-06-26] [WebCite Cache ID 6Za1O1U0s]
12. Rocco A, Heuschmann PU, Schellinger PD, Köhrmann M, Diedler J, Sykora M, et al. Glycosylated hemoglobin A1 predicts risk for symptomatic hemorrhage after thrombolysis for acute stroke. Stroke 2013 Aug;44(8):2134-2138 [FREE Full text] [doi: 10.1161/STROKEAHA.111.675918] [Medline: 23715962]
13. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000 Aug 12;321(7258):405-412 [FREE Full text] [Medline: 10938048]
14. Riddle MC, Ambrosius WT, Brillon DJ, Buse JB, Byington RP, Cohen RM, Action to Control Cardiovascular Risk in Diabetes Investigators. Epidemiologic relationships between A1C and all-cause mortality during a median 3.4-year follow-up of glycemic treatment in the ACCORD trial. Diabetes Care 2010 May;33(5):983-990 [FREE Full text] [doi: 10.2337/dc09-1278] [Medline: 20427682]
15. Krebs JD, Bell D, Hall R, Parry-Strong A, Docherty PD, Clarke K, et al. Improvements in glucose metabolism and insulin sensitivity with a low-carbohydrate diet in obese patients with type 2 diabetes. J Am Coll Nutr 2013;32(1):11-17. [doi: 10.1080/07315724.2013.767630] [Medline: 24015695]
16. Sigal RJ, Kenny GP, Wasserman DH, Castaneda-Sceppa C, White RD. Physical activity/exercise and type 2 diabetes: a consensus statement from the American Diabetes Association. Diabetes Care 2006 Jun;29(6):1433-1438. [doi: 10.2337/dc06-9910] [Medline: 16732040]
17. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, et al. Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 2012 Jun;35(6):1364-1379 [FREE Full text] [doi: 10.2337/dc12-0413] [Medline: 22517736]
18. Nundy S, Dick JJ, Chou C, Nocon RS, Chin MH, Peek ME. Mobile phone diabetes project led to improved glycemic control and net savings for Chicago plan participants. Health Aff 2014 Feb;33(2):265-272 [FREE Full text] [doi: 10.1377/hlthaff.2013.0589] [Medline: 24493770]
19. Pludwinski S, Ahmad F, Wayne N, Ritvo P. Participant experiences in a smartphone-based health coaching intervention for type 2 diabetes: A qualitative inquiry. J Telemed Telecare 2015 Jul 21:1-7. [doi: 10.1177/1357633X15595178] [Medline: 26199275]
20. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
21. NexJ Connected Wellness. URL: http://www.connectedwellness.com/ [accessed 2015-08-25] [WebCite Cache ID 6b3Lhk0n7]
22. Lenters-Westra E, Slingerland RJ. Six of eight hemoglobin A1c point-of-care instruments do not meet the general accepted analytical performance criteria. Clin Chem 2010 Jan;56(1):44-52 [FREE Full text] [doi: 10.1373/clinchem.2009.130641] [Medline: 19926777]
23. Willard-Grace R, DeVore D, Chen EH, Hessler D, Bodenheimer T, Thom DH. The effectiveness of medical assistant health coaching for low-income patients with uncontrolled diabetes, hypertension, and hyperlipidemia: protocol for a randomized controlled trial and baseline characteristics of the study population. BMC Fam Pract 2013;14:27 [FREE Full text] [doi: 10.1186/1471-2296-14-27] [Medline: 23433349]
24. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
25. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes 2003;1:29 [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
26. Crawford JR, Henry JD. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. Br J Clin Psychol 2004 Sep;43(Pt 3):245-265. [doi: 10.1348/0144665031752934] [Medline: 15333231]
27. Cheak-Zamora NC, Wyrwich KW, McBride TD. Reliability and validity of the SF-12v2 in the medical expenditure panel survey. Qual Life Res 2009 Aug;18(6):727-735. [doi: 10.1007/s11136-009-9483-1] [Medline: 19424821]
28. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2015-08-25] [WebCite Cache ID 6b3N0GIKf]
29. CANSIM (database). Ottawa, ON: Income Statistics Division, Statistics Canada; 2013 Jun 26. Table 202-0808 - Low income measures by income source and household size, 2011 constant dollars, annual (dollars) URL: http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=2020808 [WebCite Cache ID 6bdh45rLX]
30. Alegría M, Bijl RV, Lin E, Walters EE, Kessler RC. Income differences in persons seeking outpatient treatment for mental disorders: a comparison of the United States with Ontario and The Netherlands. Arch Gen Psychiatry 2000 Apr;57(4):383-391. [Medline: 10768701]
31. Steele LS, Glazier RH, Lin E. Inequity in mental health care under Canadian universal health coverage. Psychiatr Serv 2006 Mar;57(3):317-324. [doi: 10.1176/appi.ps.57.3.317] [Medline: 16524988]

32. Martinez-Millana A, Fico G, Fernández-Llatas C, Traver V. Performance assessment of a closed-loop system for diabetes management. Med Biol Eng Comput 2015 Feb 11:1-9. [doi: 10.1007/s11517-015-1245-3] [Medline: 25667016]
33. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change. 3rd edition. New York, NY: The Guilford Press; 2013.

Abbreviations

ANOVA: analysis of variance
BCCHC: Black Creek Community Health Centre
BMI: body mass index
CSEP: Canadian Society for Exercise Physiology
CWP: Connected Wellness Platform
EEP: Exercise Education Program
HADS: Hospital Anxiety and Depression Scale
HbA1c: glycated hemoglobin/hemoglobin A1c
HC: health coach
HRQOL: health-related quality of life
LOCF: last observation carried forward
NYFHT: North York Family Health Team
OHIP: Ontario Health Insurance Program
PANAS: Positive and Negative Affect Schedule
RCT: randomized controlled trial
SES: socioeconomic status
SF-12: Short Form Health Survey-12
T2DM: type 2 diabetes mellitus

Edited by P Bamidis; submitted 26.06.15; peer-reviewed by A Martinez-Millana, MS Iyengar; comments to author 28.07.15; revised version received 06.08.15; accepted 12.08.15; published 05.10.15.

Please cite as:
Wayne N, Perez DF, Kaplan DM, Ritvo P
Health Coaching Reduces HbA1c in Type 2 Diabetic Patients From a Lower-Socioeconomic Status Community: A Randomized Controlled Trial
J Med Internet Res 2015;17(10):e224
URL: http://www.jmir.org/2015/10/e224/
doi:10.2196/jmir.4871
PMID:26441467

Original Paper

Parent Engagement With a Telehealth-Based Parent-Mediated Intervention Program for Children With Autism Spectrum Disorders: Predictors of Program Use and Parent Outcomes

Brooke Ingersoll[1], PhD(Psych); Natalie I Berger[1], MA

Department of Psychology, Michigan State University, East Lansing, MI, United States

Corresponding Author:
Brooke Ingersoll, PhD(Psych)
Department of Psychology
Michigan State University
316 Physics Rd
Michigan State University
East Lansing, MI, 48824
United States
Phone: 1 517 432 8412
Fax: 1 517 353 1652
Email: ingers19@msu.edu

Related Articles:

This is a corrected version. See correction statement: http://www.jmir.org/2015/11/e239

This is a corrected version. See correction statement: http://www.jmir.org/2015/11/e257

Abstract

Background: There has been growing interest in using telehealth to increase access to parent-mediated interventions for children with ASD. However, little is known about how parents engage with such programs.

Objective: This paper presents program engagement data from a pilot study comparing self-directed and therapist-assisted versions of a novel telehealth-based parent-mediated intervention for young children with autism spectrum disorders (ASD).

Methods: Parents of young children with ASD were randomly assigned to receive a self-directed or therapist-assisted version of ImPACT Online. Parent engagement and satisfaction with the different components of the program website were examined using the program's automated data collection and a post-treatment evaluation survey. We examined the relationship between program engagement and changes in parent knowledge and implementation and participant characteristics associated with program engagement.

Results: Of the 27 parent participants, the majority were female (26/27, 96%), married (22/27, 81%), with a college degree or higher (15/27, 56%), and less than half were not employed outside of the home (10/27, 37%). The mean chronological age of the child participants was 43.26 months, and the majority were male (19/27, 70%) and white (21/27, 78%). Most of the families (19/27, 70%) resided in a rural or medically underserved area. Parents logged into the website an average of 46.85 times, spent an average of 964.70 minutes on the site, and completed an average of 90.17% of the lesson learning activities. Participants in the therapist-assisted group were more likely to engage with the website than those in the self-directed group: $F_{2,24}$=17.65, P<.001. In total, 85% of participants completed the program, with a significantly greater completion rate in the therapist-assisted group (N=27): χ^2_1=5.06, P=.03. Lesson learning activities were visited significantly more often than the supplemental activities (all Ps<.05). Multiple regression controlling for pretreatment performance indicated that program completion (beta=.51, P=.02) predicted post-treatment intervention knowledge, and program completion (beta=.43, P=.03) and group assignment (beta=-.37, P=.045) predicted post-treatment intervention fidelity. Partial correlations indicated that parent depressive symptoms at pretreatment were negatively associated with program completion (r=-.40, P=.04), but other key parent and child demographic factors were not. Post-treatment measures of website usability (r=.65, P<.001), treatment acceptability (r=.58, P=.002), and overall satisfaction (r=.58, P=.002) were all related to program completion.

Conclusions: Parent engagement and satisfaction with ImPACT Online was high for both self-directed and therapist-assisted versions of the program, although therapist assistance increased engagement. Program completion was associated with parent outcomes, providing support for the role of the website in parent learning. This program has the potential to increase access to parent-mediated intervention for families of children with ASD.

(J Med Internet Res 2015;17(10):e227) doi:10.2196/jmir.4913

KEYWORDS

autism; telehealth; parent training

Introduction

The past 20 years have seen a dramatic increase in the proliferation of telehealth—the delivery of health and mental health information via the Internet and related technologies. There are numerous potential advantages of the use of telehealth to augment or even replace traditional service delivery models. Perhaps most appealing is its potential to increase access to evidence-based interventions for individuals living in rural and underserved areas at a significantly reduced cost [1]. Such programs have been shown to greatly improve care for patients with chronic diseases, such as diabetes, heart disease, and asthma [2-4], and increase access to evidence-based health promotion [5-7], psychological [8-10], and parenting interventions [11]. Patients are often very satisfied with the care they receive through telehealth services [12,13], and both efficacy and effectiveness trials have found moderate to large effects of telehealth interventions on participant knowledge and behavior [14-16].

Telehealth's ability to increase access to services in a cost effective manner [1] indicates that it may be an attractive option for the treatment of autism spectrum disorders (ASDs). Given that children with ASD experience significant difficulties accessing traditional services [17] and that the prevalence rate of ASD has increased significantly in the past 15 years [18], alternative service delivery models are necessary in order to provide this unique population with sufficient treatment options. Indeed, survey-based studies have indicated that approximately one third of children with ASD have problems obtaining specialty care [19], and approximately 40% experience difficulties accessing desired services more generally [20]. Difficulty accessing services is compounded for children with ASD from underserved areas [21] and those belonging to racial and ethnic minority families [17,20].

There has been growing interest in the use of telehealth for delivering parent-mediated interventions for children with ASDs [22-26]. Children with ASD experience pervasive deficits in social communication and the presence of restricted and repetitive behaviors [27], which can significantly impact the quality of life of the child and family [28-30]. Parent-mediated intervention, in which the parent is trained to use intervention techniques directly with their child, has been widely recognized as a critical component of effective autism intervention [31,32]. Research indicates that parents can learn to use evidence-based intervention techniques with fidelity [33-36], and parent use of the techniques results in improvements in child skills and behavior [34,37]. Further, parent-mediated intervention is associated with improvements in family quality of life [38,39].

Despite its benefits, parent-mediated intervention is highly under-utilized in community settings [40]. Barriers include a shortage of trained professionals, lengthy waitlists, limited financial resources and transportation, lack of child care, geographic isolation, and time limitations [41-43], many of which are compounded in families who live in rural and underserved areas [44].

Although the evidence base at this point is very limited, several quasi-experimental and single-case design studies have demonstrated the potential of telehealth to deliver parent-mediated intervention for children with ASD. Parents generally find such programs to be acceptable [26], and program use is associated with gains in parent knowledge [45-48] and use of evidence-based intervention procedures [26,36,49]. Importantly, there is also emerging evidence for improvement in child social communication skills as a result of parent use of such programs [23,36,49].

These studies are encouraging and suggest that telehealth, should it prove efficacious in larger trials, may be an effective method for increasing access to parent-mediated interventions for families of children with ASD. However, there is very little known about factors that influence parent engagement with such programs. For example, it is not clear what types of parents choose to enroll in telehealth-based parent-mediated intervention programs and whether they are different from parents who enroll in traditional parent-mediated intervention programs. It is also unclear whether specific supports provided during the program influence program engagement. Several studies have suggested that therapist assistance can positively influence participant engagement with cognitive-behavioral therapy (CBT)-based telehealth interventions for mood and anxiety [10], and remote coaching by a therapist after participation in a self-directed, telehealth-based parent-mediated intervention was shown to improve parents' ability to implement an imitation intervention with their child with ASD [26]. However, the role of therapist assistance in promoting program engagement with telehealth-based parent-mediated intervention programs for ASD is unknown. Further, although research has indicated that parent participation in such programs is associated with gains in parent knowledge and skill use [23,50], additional research is needed to examine to what extent parent engagement with the website used to deliver programs is related to these important parent outcomes.

Finally, little is known about parent and/or child characteristics that are associated with program engagement in telehealth-based parent-mediated intervention. Across a range of health information technologies, participants' computer/Internet self-efficacy, as well as the perceived usefulness and ease of

use of the technology have been associated with technological acceptance [51], suggesting that these factors may influence parents' participation in telehealth-based parent-mediated intervention programs. However, the majority of these studies examined behavioral intention or self-reported use, rather than actual use or engagement with the technology, and studies that have used objective measures of program use have sometimes failed to find a relationship with these variables [52,53]. Treatment acceptability, or the degree to which consumers view an intervention as appropriate, fair, and reasonable for the presenting problem [54], has been associated with client engagement in both behavioral and medical treatments [55,56]; however, little research has examined the relationship between treatment acceptability and engagement in telehealth-based interventions [57]. Sociodemographic factors, such as age, gender, race/ethnicity, marital status, education, and employment, have often failed to emerge as consistent predictors of engagement in telehealth programs more broadly [51]. At the same time, these factors have been found to influence engagement in traditional parent-mediated intervention programs, with lower socioeconomic status and maternal mental health problems being particularly salient predictors of dropout [58]. It is not yet known whether these same characteristics apply to telehealth-based parent-mediated intervention for children with ASD. This is a particularly important area to examine, given that lower socioeconomic status is associated with significantly poorer access to services for children with ASD [17,21].

In this paper, we present data about program engagement from a pilot study of ImPACT Online, a telehealth-based parent-mediated intervention for young children with ASD. Specifically, we examined (1) the demographics of parent and child participants who enrolled in the program, (2) the overall level of program engagement among parents of children with ASD, the components of the website that parents used most and least often, and parents' perceptions of the usefulness of the different program components, (3) whether therapist assistance influenced engagement with or perceived usefulness of the program, (4) whether parent engagement with the website was associated with changes in parent intervention knowledge and intervention fidelity, and (5) whether pretreatment participant characteristics (parent computer/Internet fluency, parent depressive symptoms, family demographics, child severity) and/or parents' experience with the program (website usability, treatment acceptability, overall satisfaction) were associated with greater engagement with the website.

Methods

Participants

Participants included 28 parents of a child with ASD between the ages of 27 and 73 months. Participating families were recruited from agencies serving children with ASD (eg, early intervention programs, diagnostic centers, parent support groups) and online. Recruitment focused on underserved communities as defined by residence in a rural (based on the United States Department of Agriculture Rural Development designation) and/or medically underserved area/provider shortage area (based on the Health Resources and Services Administration medically underserved area and population designation). All children met criteria for Autistic Disorder or Pervasive Developmental Disorder, Not Otherwise Specified based on the *Diagnostic and Statistical Manual – 4th Edition, text revision* (DSM-IV-TR) criteria [59] and the Autism Diagnostic Observation Schedule, generic or 2nd edition (ADOS-G/ADOS-2) [60]. Parents had to be proficient in English, although other languages could be spoken in the home. For two-parent households, both parents were able to use the online tutorial if they wanted; however, only 1 parent participated in data collection and received online coaching (for the therapist-assisted group).

Families without a personal computer, webcam, or high-speed Internet in the home were provided the required technology. One family suspended treatment for 7 months in the middle of the program due to a significant health problem. This family's data were excluded from analysis as their pre-post data were not comparable to the other families', yielding a total of 27 participants. All parents gave informed consent for their own and their child's participation under the oversight of Michigan State University's Human Research Protections Program. Table 1 presents participant characteristics by group.

Table 1. Participant demographic information.

	Overall	Group		Test statistic	*P* value
		Self-directed	Therapist-assisted		
Parent characteristics					
Gender (% female)	96	92	100	1.12[a]	.29[a]
Education level (% college degree)	56	46	64	0.90[a]	.34[a]
Marital status (% married)	81	92	71	1.95[a]	.16[a]
Employment status (% not employed)	37	46	29	0.65[a]	.42[a]
Residence in underserved area, %	70	77	64	0.52[a]	.47[a]
Computer/Internet self-efficacy (CEWFS)	35.57 (4.25)	35.09 (4.76)	36.00 (3.89)	0.50[b]	.62[b]
Depressive symptoms (CES-D)	10.89 (8.26)	10.23 (6.98)	11.50 (9.52)	0.39[b]	.70[b]
Child characteristics					
Gender (% male)	70	61	79	0.94[a]	.33[a]
Race/Ethnicity (% white)	78	92	64	3.06[a]	.08[a]
Chronological age in months, mean (SD)	43.26 (12.58)	46.08 (13.18)	41.57 (12.24)	-0.71[b]	.48[b]
Nonverbal mental age in months, mean (SD)	24.83 (11.57)	25.42 (13.92)	24.29 (9.38)	-0.25[b]	.80[b]
Verbal mental age in months, mean (SD)	20.44 (10.11)	19.15 (9.63)	21.64 (10.74)	0.63[b]	.53[b]
ADOS comparison score, mean (SD)	7.00 (1.61)	7.00 (1.87)	7.00 (1.35)	0.00	.99[b]
Nonstudy intervention hrs/wk, mean (SD)	12.98 (10.15)	13.62 (10.96)	12.38 (9.70)	-0.31[b]	.76[b]

[a]Chi square.

[b]*t* test.

Design Overview

Participating families provided demographic information and completed a standardized battery of assessments of parent and child functioning in the lab and family home at intake. Children were matched on expressive language on the Mullen Scales of Early Learning [61] and then randomly assigned to the self-directed or therapist-assisted group using a coin flip. At post-treatment, parents completed measures of parent learning, a survey-based evaluation of ImPACT Online, along with other measures of parent and child functioning (not presented here).

Interventions

Self-Directed Group

Parents assigned to the self-directed group received access to the secure, password-protected ImPACT Online website for up to 6 months. The program was adapted from Project ImPACT, which is a naturalistic, developmental-behavioral, parent-mediated intervention for young children with ASD [62]. The website contained 12 self-directed lessons, each of which took approximately 80 minutes to complete. Parents were encouraged to complete one lesson per week and to practice the intervention techniques with their child between each lesson. Each lesson included six learning activities that were designed to be completed in order. The topic was introduced via a Narrated Slideshow with embedded video clips of an expert therapist or parent walking step-by-step through each technique and providing tips for successful implementation. The same information was elaborated on in a Written Manual that could be printed for later reference. After each slideshow, parents completed a Self-Check Quiz and short Interactive Exercises, in which they viewed short video clips of adult-child interactions and were asked to identify accurate use of the intervention techniques. Parents were provided with immediate feedback on their performance to facilitate learning. At the end of the lesson, parents completed a Homework Plan in which they identified child goals, activities in which they would practice, and how they would implement the specific techniques within those activities. After completing the homework plan, parents responded to Reflection Questions that asked them to report how they used the techniques, their child's response, what aspects went well, and what aspects were challenging.

Parents could also access supplemental material outside of the weekly lessons. Supplemental material included a Video Library that contained longer video examples of adults using all of the intervention techniques together, a Resources page with links to autism informational websites, and a moderated Forum that allowed parents to share information with other parents. Parents received weekly "Tip of the Week" emails that provided tips for implementing the intervention techniques along with a link to the program to encourage program use. Parents in the self-directed group were able to contact project staff via phone or email for assistance with technology-related problems (eg, difficulty with logins, problems playing video) but were given

no assistance or support in learning the intervention from project staff outside of the self-directed website. See Multimedia Appendix 1 for screenshots.

Therapist-Assisted Group

Parents assigned to the therapist-assisted group were given access to the ImPACT Online website and were encouraged to work through the program at the same pace as the self-directed group. Parents also received two 30-minute remote coaching sessions per week (24 total sessions) via Skype video conferencing software by a trained therapist to assist them in learning the intervention. Skype was selected as the primary video conferencing software during the development phase of the program based on feedback from focus groups with parents and providers indicating that parents would be most comfortable using Skype (over other available systems) based on its simplicity and familiarity. Parents in this study were made aware that Skype was not Health Insurance Portability and Accountability Act (HIPAA)-compliant during the consent process.

The first coaching session of the week involved the coach and parent and was used to help clarify the content of the relevant lesson and help the parent apply the information to their child. The second coaching session of the week involved the coach, parent, and child and was used to provide the parent with "live" feedback on their use of the intervention techniques as they practiced with their child.

Measures

Participant Characteristics

Parents provided demographic information about themselves and their child at intake. Parent demographic variables included gender, marital status, education level, employment status, and residence in a rural and/or medical professional shortage area. Child demographic variables included age, gender, and race/ethnicity. Parents also provided information on the type and hours per week of all nonstudy treatments (eg, speech-language therapy, therapeutic preschool) their child received. Children were administered the Mullen Scales of Early Learning [61] by the study team to determine their developmental age and the ADOS-G or ADOS-2 [60] to determine autism severity.

Computer and Internet Fluency

Parents completed a modified version of the Computer-Email-Web Fluency Scale (CEWFS) [63] at intake to assess their comfort with computers and Internet technology. Items such as (1) How frequently do you conduct a search using an Internet search engine? and (2) How frequently do you send or receive email? were rated on a 5-point scale from never (1) to daily (5). In addition, parents were asked to indicate on a 7-point scale how many hours per week they used the Internet. Scores could range from 8-42, with higher scores indicative of greater fluency. Cronbach alpha for the modified scale was .70, indicating adequate internal consistency. Data were missing on the CEWFS for 4 participants.

Depressive Symptoms

Parents completed the Center for Epidemiological Studies-Depression Scale (CES-D) [64] at intake as a measure of depressive symptoms. Respondents used a 4-point scale to rate how often they experienced 20 different symptoms of depression over the past week, with higher scores indicating a greater degree of depressive symptomatology.

Intervention Knowledge

Parents completed the ImPACT Knowledge Quiz. This 20-item, multiple-choice quiz measures comprehension of the key elements of Project ImPACT, at intake and post-treatment to measure changes in their intervention knowledge. Data were missing at post-treatment for one participant.

Intervention Fidelity

Parents were videotaped during a parent-child interaction in their home at intake and post-treatment to measure changes in their intervention fidelity. Parents were asked to play with their child for 10 minutes and have a small snack or meal with their child. Parent behavior was scored for correct use of the intervention strategies using the Project ImPACT intervention fidelity checklist [62]. Each of the six fidelity dimensions was scored on a scale of 1 (Parent does not implement throughout session) to 5 (Parent implements throughout session) and then averaged to form an overall fidelity rating for each routine. The fidelity dimensions included (1) Focus on your child, (2) Adjust your communication, (3) Create opportunities for engagement, (4) Teach your child new communication skills, (5) Teach your child new play skills, and (6) Pace the interaction. Ratings for the play and snack routines were averaged to form an overall fidelity rating for each time point. Data were missing at post-treatment for one participant.

Program Engagement

Program engagement was calculated from the ImPACT Online website's electronic tracking of user behavior. Several different metrics of program engagement were calculated, including (1) average number of logins to the site, (2) average duration of time spent on the site across the intervention period, and (3) program completion, which was defined as completion of at least 75% of total learning activities across the 12 lessons (out of a possible 71). In addition, we examined the average duration of each login, the average number of days to program completion, and the times of day when logins occurred.

To better understand how parents used the program, we calculated the average number of visits made to each of the lesson learning activities (slideshow, manual, self-check, exercises, homework, reflection) and supplemental materials (video library, resources, forum), as well as the percent of each type of lesson learning activity completed.

Program Evaluation

At post-treatment, parents were asked to complete an evaluation survey measuring treatment acceptability, website usability, and overall program satisfaction. All items were rated using a 7-point Likert scale, with higher scores indicative of greater satisfaction. Parents rated the acceptability of the treatment using a 30-item questionnaire [65] adapted from the Treatment

Evaluation Inventory [66] and the Behavior Intervention Rating Scale [67]. Cronbach alpha for the treatment acceptability measure was .93. Parents rated the usability of the website using an 11-item questionnaire developed for this project. For this measure, parents rated the perceived helpfulness of each program component using the following statement: "The [individual program component]s were helpful for learning the ImPACT Online intervention", as well as 3 additional items: (1) "It was easy to find information on the program website," (2) "I used the website to complete the skills-training sessions," and (3) "I understood the audio and text information that was presented." Cronbach alpha for the website usability scale was .81. The average of two additional items, (1) "I used the intervention with my child regularly" and (2) "I would recommend this program to other parents of young children with social-communication difficulties, was used as a measure of overall program satisfaction."

Parents in the coaching group were also asked to rate the quality of the therapist relationship using the following items: (1) "My coach was interested in me," (2) "My coach understood me," and (3) "My coach understood my child, and the perceived helpfulness of (i) the weekly discussion sessions with the trainer, (ii) the weekly coaching sessions with my child."

Results

The data were examined for normality using visual analysis of the data distributions and the inspection of the skewness coefficient. All variables were found to be normally distributed with the exception of overall satisfaction, which was subject to ceiling effects. Thus, we used nonparametric tests (Mann-Whitney U, Spearman rank correlation) for analyses involving the overall satisfaction variable and parametric tests for all other analyses.

Participant Characteristics

As shown in Table 1, the majority of parent participants were female (26/27, 96%), married (22/27, 81%), with a college degree or higher (18/27, 66%). Less than half of parents were not employed outside of the home (10/27, 37%). The majority of child participants were male (19/27, 70%) and white (19/27, 78%). The children's mean chronological age was 43.26 months. All children exhibited a significant developmental delay, with an average nonverbal mental age of 24.83 and verbal mental age of 20.44, and their average autism severity fell in high moderate range. These characteristics are relatively typical of families who participate in parent-mediated intervention studies for young children with ASD [68]. Consistent with our effort to recruit families in underserved areas, 70% (21/27) of our sample resided in a rural or medical underserved area. Two parents requested technology (2 needed a computer with webcam and 1 needed high-speed Internet) in order to be able to complete the program in their home. Scores on the CEWFS ranged from 26 to 40, with average scores toward the upper end of the scale (35.8), suggesting that most parents felt fluent with computer and Internet technology. The majority of parents reported using computers 10 or more hours per week, although 22% reported using computers 4 hours a week or less. Parents reported that their children received a variety of different intervention services (eg, special education, applied behavior analysis, speech therapy, occupational therapy, play groups), and the number of hours of nonstudy intervention they received per week ranged from 0.5 to 38.5 hours. We used independent sample *t* tests and chi-square tests as appropriate to examine group difference on pretreatment characteristics. There were no significant between-group differences on participant characteristics at intake.

Program Engagement

As can be seen in Table 2, there was a high rate of program engagement as measured by the various metrics. Parents logged into the website an average of 46.85 times and spent an average of 964.70 minutes on the site over the intervention period. Eighty-five percent of parents (23/27) completed the program—a completion rate that is similar to the traditional, in-person version of the program [69]. Parents averaged 21.23 minutes per login and took an average of 133.87 days to complete the program. Parent logins were equally likely to occur throughout the day (8-11 a.m.: 19.67%; 11 a.m.-2 p.m.: 15.90%; 2-5 p.m.: 19.28%) and evening (5-8 p.m.: 17.52%; 8-11 p.m.: 19.70%)]. Parents logins were significantly less likely in the early morning (5-8 a.m.: 4.45%) and overnight (11 p.m.-5 a.m.: 3.48%). It should be noted that almost half of all logins (45.19%) occurred between 5 p.m. and 8 a.m., which are times of day when traditional therapy is typically not available.

We examined group differences on the two continuous metrics of program engagement (number of logins, duration on site) using a multivariate analysis of variance (MANOVA), and program completion using chi-square. There was a significant effect of group on program engagement: $F_{2,24}$=17.65, P<.001. Follow-up tests revealed that the therapist-assisted group demonstrated significantly greater program engagement on both metrics than the self-directed group (see Table 2). The therapist-assisted group was also significantly more likely to complete the program than the self-directed group (N=27): χ^2_1=5.06, P=.03.

Table 2. Program engagement.

	Overall, mean (SD)	Self-directed, mean (SD)	Therapist-assisted, mean (SD)	Test statistic	*P* value
Number of logins	46.85 (22.30)	29.54 (13.76)	62.93 (15.55)	34.68	<.001
Duration of time on site	964.70 (518.49)	707.04 (402.41)	1203.94 (510.06)	7.81	.01
Participants completing program, %	85	69	100	5.06	.03
Days to completion	133.87 (38.30)	148.56 (41.17)	124.42 (34.55)	-1.52	.14
Duration per login	21.23 (6.70)	23.38 (4.91)	19.23 (7.66)	-1.66	.11
Visits to lesson learning activities					
1. Slideshows	18.22 (6.50)[a]	16.54 (4.99)	19.79 (7.48)	1.32	.20
2. Manual	23.59 (11.40)[a]	21.15 (9.02)	25.86 (13.17)	1.07	.29
3. Self-check	14.56 (4.61)[b]	12.69 (4.25)	16.29 (4.38)	2.16	.04
4. Exercises	14.15 (5.11)[b]	11.62 (4.63)	16.50 (4.50)	2.79	.01
5. Homework	31.93 (19.54)[c]	22.46 (10.85)	40.71 (21.95)	2.70	.01
6. Reflection	20.78 (9.17)[a]	16.69 (9.69)	24.57 (7.02)	2.43	.02
Visits to supplemental materials					
7. Video library	4.04 (3.45)[d]	2.23 (1.79)	5.71 (3.81)	3.00	.01
8. Resources	1.85 (2.32)	1.69 (2.90)	2.00 (1.71)	0.34	.74
9. Forum	2.33 (1.82)[d]	2.54 (2.03)	2.14 (1.66)	-0.56	.58

[a]Significantly greater than 3, 4, 7, 8, 9.

[b]Significantly greater than 7, 8, 9.

[c]Significantly greater than 1, 2, 3, 4, 6, 7, 8, 9.

[d]Significantly greater than 8.

We then compared the average number of visits parents made to each program component (ie, learning activities and supplemental materials) using a mixed-model, repeated-measures ANOVA with group as a between-subjects variable and program component as a within-subjects variable. There was a main effect of group ($F_{1,25}$=241.86, P<.001) favoring the therapist-assisted group, and a main effect of program component ($F_{8,200}$=57.66, P<.001). Follow-up paired *t* tests indicated that parents were most likely to visit the homework, followed by the reflection, manual, and slideshow, and then the self-check and exercises. Parents visited each of the lesson learning activities more often than the supplemental materials (all Ps≤.05). There was also a significant group x program component interaction. Follow-up independent samples *t* tests indicated that the therapist-assisted group visited the self-check, exercises, homework, reflection, and video library significantly more than the self-directed group (all Ps<.05). However, the groups did not differ in number of visits to the slideshow, manual, resources, and forum.

The number of visits to each learning activity contained in the lessons may be slightly misleading as a metric of engagement, as the different activities may pull for a different number of visits. For example, parents may only visit the self-check questions once in order to assess their comprehension, whereas they might return several times to the homework page in order to complete or review their homework plan. Therefore, we also examined the percent of learning activities that parents completed, regardless of the number of visits that this took. Again, there was a main effect of group ($F_{1,25}$=4.37, P=.047) favoring the therapist-assisted group, and a main effect of learning activity ($F_{5,125}$=3.33, P=.007). But there was not a group x learning activity interaction (P>.05). Follow-up paired *t* tests indicated that parents completed the reflection (mean 83.8%, SD 5.70%) significantly less often than each of the other lesson learning activities (range 87.65-92.90%; all Ps<.05), with the exception of the slideshow for which the difference was marginal (P=.095). There was no difference in completion rates between any of the other learning activities.

Program Evaluation

Parents rated the intervention content as highly acceptable (mean 6.07, SD 0.79) and the website as highly usable (mean 6.36, SD 0.57). In addition, parents indicated that they were highly satisfied with the program overall (mean 6.56, SD 0.71). The therapist-assisted group rated the quality of the therapist relationship uniformly highly (mean 6.99, SD 0.06). There were no group differences on ratings of treatment acceptability (t_{25}=1.79, P=.09) or website usability (t_{25}=1.28, P=.21). There was a marginally significant group difference on parents' rating of overall program satisfaction, with parents in the therapist-assisted group rating the program significantly higher

(mean 6.86, SD 0.23) than those in the self-directed group (mean 6.23, SD 0.90; Z=1.92, P=.054).

Effect of Program Engagement on Parent Outcomes

Next, we examined the degree to which program engagement was predictive of parent intervention knowledge and intervention fidelity at post-treatment. We chose to focus on program completion (ie, completion of at least 75% of all lesson learning activities) because it is a conservative indicator of exposure to program content across intervention delivery modalities [70]. For these analyses, we employed a series of linear regressions, with parents' intervention knowledge or intervention fidelity at post-treatment as the dependent variable, and parents' pretreatment score for the relevant measure, group assignment, and program completion as independent variables.

The full model predicting parents' post-treatment intervention knowledge was statistically significant: $F_{3,25}$=4.81, P=.01. After controlling for pretreatment intervention knowledge, program completion was a significant predictor of post-treatment intervention knowledge: beta=.45, t=2.45, P=.02 (Table 3). This finding suggests that regardless of group assignment, parents who completed the program experienced more improvement in their intervention knowledge from pre- to post-treatment. The full model predicting parents' post-treatment intervention fidelity was also significant: $F_{3,25}$=5.87, P=.004. After controlling for pretreatment intervention fidelity, both group assignment (beta=-.37, t=-2.12, P=.045) and program completion (beta=.43, t=2.30, P=.03) were significant predictors of intervention fidelity at post-treatment. This result indicates that both therapist assistance and program completion make independent contributions to gains in parents' accurate use of the intervention techniques with their child.

Table 3. Predictors of parent intervention knowledge and parent intervention fidelity at post.

	Post-treatment performance			
	Parent knowledge		Parent fidelity	
Predictors	Beta	*t*	Beta	*t*
Pretreatment performance	0.40	2.39[a]	0.38	2.21[a]
Group assignment	0.04	0.21	-0.37	-2.12[a]
Program completion	0.51	2.81[a]	0.43	2.30[a]

[a]P<.05.

Predictors of Program Engagement

Finally, we examined the extent to which parent and child demographic factors at pretreatment and parent program evaluation variables at post-treatment were associated with program engagement. Again, we focused on program completion as our measure of program engagement. To do this, we calculated partial correlations between participant demographic variables at intake, program evaluation variables at post-treatment, and program completion, controlling for group assignment. Parent demographic variables included education level (no college degree vs college degree or higher), marital status (married vs not married), employment status (employed full or part time vs not employed), residence in underserved area, computer/Internet self-efficacy, and depressive symptoms. We did not examine parent gender as there was only 1 male parent participant. Child demographic variables included age, gender, race/ethnicity (majority vs minority), developmental quotient, autism severity (ADOS comparison score), and hours of nonstudy intervention. Program evaluation measures included treatment acceptability, website usability, and overall program satisfaction.

Of all of the pretreatment parent and child demographic factors, only parent depression was significantly associated with program completion: r_{24}=-.42, P=.04 (Table 4). Parents who rated themselves higher in depressive symptoms at pretreatment were significantly less likely to complete the program. All program evaluation variables (treatment acceptability: r_{24}=.58, P=.002; website usability: r_{24}=.65, P<.001; and overall program satisfaction: r_{24}=.58, P=.002) were significantly associated with program completion.

Table 4. Partial correlates of program completion controlling for group assignment.

	Program completion
Parent	
Education level	0.20
Marital status	-0.10
Employment status	0.04
Residence in underserved area	-0.24
Computer/internet self-efficacy	0.16
Depressive symptoms	-0.40[a]
Child	
Age	0.13
Gender	-0.12
Race/Ethnicity	0.09
Developmental quotient	0.10
Autism severity	-0.15
Hours of nonstudy intervention	0.24
Website usability	0.65[c]
Treatment acceptability	0.58[b]
Overall program satisfaction	0.58[b]

[a]$P<.05$.

[b]$P<.01$.

[c]$P<.001$.

Discussion

Principal Findings

In this study, we examined parent engagement with a novel, telehealth-based parent-mediated intervention for children with ASD. Overall findings indicated that parent engagement and satisfaction with ImPACT Online was high for both self-directed and therapist-assisted versions of the program, although therapist assistance increased engagement. Parent outcomes were associated with program completion, providing support for the role of the website in parent learning.

Parents who enrolled in this study were quite similar to participants in studies of traditional parent-mediated interventions for children with ASD [68]. Like most of these studies, the majority of our parent participants were married, college-educated mothers who were not employed outside of the home, and our child participants were white males, with a moderate ASD and significant developmental delay. However, unlike many previous studies, the majority of our sample lived in a rural or medically underserved area or were a medically underserved population. This suggests that telehealth-based parent-mediated intervention appeals to similar families who seek out traditional parent-mediated intervention (at least those who are involved in research) and may be able increase access to families in underserved areas. At the same time, there were few very low socioeconomic status families in our sample, and thus participation was not representative of the range of families who could potentially benefit from this type of intervention. Thus, future research should examine the potential reach of telehealth-based parent-mediated intervention in the population of families of young children with ASD as well as the representativeness of parents who enroll in these interventions.

We found high rates of program engagement among parents, with overall completion rates similar to those found in studies of traditional parent-mediated intervention programs for children with ASD [71,72]. Parents also rated the program very highly in terms of the acceptability of the treatment, the usability of the website, overall satisfaction with the program, and the quality of the therapist relationship (for the therapist-assisted group). Many of the parents engaged with the website during times of day when traditional parent training would not occur (eg, early mornings, evenings, weekends). Our data suggest that placing instructional activities online does not compromise parent engagement or parent satisfaction with parent-mediated intervention and may allow families to learn during times that are more convenient, such as before or after work and when children are in bed.

Parents were more likely to engage with certain program components than others. Specifically, the learning activities contained in the lessons were visited significantly more often than supplemental materials. Among these learning activities, parents were significantly less likely to complete the reflection than the other learning activities. Interestingly, the forum was rarely visited and only three parents posted to the forum even

once. Models of technological acceptance that claim the perceived usefulness and ease of use a technology (or in our case, component) have a strong influence on its use [73]. This finding suggests that adjustments could be made to those components that were used less often to increase their perceived usefulness and/or ease of use.

Therapist assistance positively influenced parent engagement with the program. This finding is consistent with research on telehealth-based behavioral interventions; therapist support increases engagement with the program as well as overall outcomes [10,74]. In our study, therapist assistance was associated with a higher rate of program completion, as well as more frequent logins and a greater duration of time spent on the site. Therapist assistance appeared to be particularly beneficial for encouraging parents to visit specific components of the program, namely the self-check questions, exercises, homework, reflection, and video library. It may be that these components are more challenging or less motivating for parents to engage with on their own. Alternately, it may be that parents accessed these components during their coaching sessions after having completed them on their own. For example, parents were encouraged to go over their homework plan and reflection with their coach. The fact that therapist assistance did not seem to differentially affect the specific learning activities that were completed supports this possibility and suggests that a greater understanding of how therapist assistance influences parents' use of the different program component would be important. Therapist assistance did not affect parents' overall perception of acceptability of the treatment or the usability of the website. It did, however, marginally impact the parents' overall satisfaction with the program. This greater level of satisfaction may have impacted parents' willingness to engage with the website.

Parent completion of the learning activities on the website was a significant predictor of gains in parent knowledge and intervention fidelity. This finding suggests that program engagement was associated with gains in both conceptual and procedural knowledge of the intervention regardless of group assignment. Our regression model indicated that group assignment was also a significant predictor of gains in parent fidelity, which suggests that therapist assistance provided an additional benefit above and beyond completion of the tutorial for procedural knowledge. This finding is consistent with a large body of literature on parent training, which suggests that coaching is an important component for improving parent use of intervention techniques [43,49,75,76].

We examined a number of pretreatment parent and child demographic variables as potential predictors of parent program engagement. Only one of these variables, parent depressive symptoms, was significantly associated with program completion. Parents who reported more depressive symptoms at intake were less likely to complete the program than parents who reported fewer depressive symptoms. This finding is consistent with research on traditional parent training that has found that pretreatment parent psychopathology, particularly depression, increases the likelihood of attrition [58,77,78]. Thus, future studies should aim to examine the role of depression as a moderator of treatment effects for telehealth-based parent-mediated intervention. Our failure to find an association between program engagement and sociodemographic factors may have been due to our small sample size, and rather homogenous participant characteristics. However, a number of highly powered studies have failed to find an association between sociodemographic factors and use of health information technology [52,53,79]. Surprisingly, we did not find an association between program engagement and computer/Internet self-efficacy. It may be that computer/Internet self-efficacy may be more likely to affect a parent's choice to enroll in a telehealth-based parent-mediated intervention program rather than to engage with the program once enrolled. Indeed, overall levels of computer/Internet self-efficacy in this sample were relatively high, and almost all participants had access to a computer with Internet in their home prior to study enrollment. Thus, research is needed to determine uptake and engagement with telehealth-based parent-mediated intervention among parents with lower computer/Internet self-efficacy and those with limited computer and Internet access.

In contrast, all of the post-treatment evaluation items were related to program engagement. The relationship between these factors and engagement is complex. Perceived usefulness and ease of use have consistently been found to predict behavioral intention to use a new technology and are considered important antecedent determinants of technological acceptance [80,81]. At the same time, engagement with a new technology can influence its perceived usefulness and ease of use [82]. Similarly, treatment acceptability has been found to predict engagement in intervention [83]; however, it can also be affected by intervention participation [84]. As these measures were taken post-treatment in our study, the direction of influence is not clear. However, these findings highlight the importance of participants' experience with both the technology (website) and the content (treatment) of the telehealth program for engagement. It should be noted that the website usability questionnaire that we used was developed for this study. Our findings were consistent with previous work in this area. Nonetheless, this is a limitation and future work should employ standard evaluations of perceived usefulness and ease of use, such as the System Usability Scale [85].

Conclusion

Taken together, these data provide support for the potential of telehealth to deliver parent-mediated interventions to parents of children with ASD. Additional research will be important to determine whether the same patterns and predictors of program engagement emerge in larger samples.

Acknowledgments

This project was supported by Idea Development Award #W81XWH-10-1-0586 from the Department of Defense's Autism Research Program. We would like to thank the parents and children who participated in this work. We would also like to thank Nikki Bonter, MSW and the graduate and undergraduate research assistants who worked on this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of ImPACT Online.

[PDF File (Adobe PDF File), 448KB - jmir_v17i10e227_app1.pdf]

References

1. Jennett PA, Affleck HL, Hailey D, Ohinmaa A, Anderson C, Thomas R, et al. The socio-economic impact of telehealth: A systematic review. J Telemed Telecare 2003;9(6):311-320. [doi: 10.1258/135763303771005207] [Medline: 14680514]
2. McLean S, Chandler D, Nurmatov U, Liu J, Pagliari C, Car J, et al. Telehealthcare for asthma: A Cochrane review. CMAJ 2011 Aug 9;183(11):E733-E742 [FREE Full text] [doi: 10.1503/cmaj.101146] [Medline: 21746825]
3. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K. Home telehealth for diabetes management: a systematic review and meta-analysis. Diabetes Obes Metab 2009 Oct;11(10):913-930. [doi: 10.1111/j.1463-1326.2009.01057.x] [Medline: 19531058]
4. Polisena J, Tran K, Cimon K, Hutton B, McGill S, Palmer K, et al. Home telemonitoring for congestive heart failure: A systematic review and meta-analysis. J Telemed Telecare 2010;16(2):68-76. [doi: 10.1258/jtt.2009.090406] [Medline: 20008054]
5. Ammann R, Vandelanotte C, de VH, Mummery WK. Can a website-delivered computer-tailored physical activity intervention be acceptable, usable, and effective for older people? Health Educ Behav 2013 Apr;40(2):160-170. [doi: 10.1177/1090198112461791] [Medline: 23077157]
6. Carlson LE, Lounsberry JJ, Maciejewski O, Wright K, Collacutt V, Taenzer P. Telehealth-delivered group smoking cessation for rural and urban participants: Feasibility and cessation rates. Addict Behav 2012 Jan;37(1):108-114. [doi: 10.1016/j.addbeh.2011.09.011] [Medline: 21968227]
7. Haugen HA, Tran ZV, Wyatt HR, Barry MJ, Hill JO. Using telehealth to increase participation in weight maintenance programs. Obesity 2007 Dec;15(12):3067-3077. [doi: 10.1038/oby.2007.365] [Medline: 18198316]
8. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
9. Nelson E, Barnard M, Cain S. Treating childhood depression over videoconferencing. Telemed J E Health 2003;9(1):49-55. [doi: 10.1089/153056203763317648] [Medline: 12699607]
10. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
11. Wade SL, Oberjohn K, Burkhardt A, Greenberg I. Feasibility and preliminary efficacy of a web-based parenting skills program for young children with traumatic brain injury. J Head Trauma Rehabil 2009;24(4):239-247. [doi: 10.1097/HTR.0b013e3181ad6680] [Medline: 19625863]
12. Gustke SS, Balch DC, West VL, Rogers LO. Patient satisfaction with telemedicine. Telemedicine Journal 2000 May;6(1):5-13. [doi: 10.1089/107830200311806]
13. Mair F, Whitten P. Systematic review of studies of patient satisfaction with telemedicine. BMJ 2000 Jun 3;320(7248):1517-1520 [FREE Full text] [Medline: 10834899]
14. Omboni S, Guarda A. Impact of home blood pressure telemonitoring and blood pressure control: A meta-analysis of randomized controlled studies. Am J Hypertens 2011 Sep;24(9):989-998. [doi: 10.1038/ajh.2011.100] [Medline: 21654858]
15. Xie Y, Dixon JF, Yee OM, Zhang J, Chen YA, Deangelo S, et al. A study on the effectiveness of videoconferencing on teaching parent training skills to parents of children with ADHD. Telemed J E Health 2013 Mar;19(3):192-199. [doi: 10.1089/tmj.2012.0108] [Medline: 23405952]
16. 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]
17. Liptak GS, Benzoni LB, Mruzek DW, Nolan KW, Thingvoll MA, Wade CM, et al. Disparities in diagnosis and access to health services for children with autism: data from the National Survey of Children's Health. J Dev Behav Pediatr 2008 Jun;29(3):152-160. [doi: 10.1097/DBP.0b013e318165c7a0] [Medline: 18349708]

18. Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators, Centers for Disease Control and Prevention (CDC). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. MMWR Surveill Summ 2014 Mar 28;63(2):1-21 [FREE Full text] [Medline: 24670961]
19. Krauss MW, Gulley S, Sciegaj M, Wells N. Access to specialty medical care for children with mental retardation, autism, and other special health care needs. Ment Retard 2003 Oct;41(5):329-339. [doi: 10.1352/0047-6765(2003)41<329:ATSMCF>2.0.CO;2] [Medline: 12962535]
20. Kohler FW. Examining the services received by young children with autism and their families: A survey of parent responses. Focus Autism Other Dev Disabl 1999 Jan 01;14(3):150-158. [doi: 10.1177/108835769901400304]
21. Ruble LA, Heflinger CA, Renfrew JW, Saunders RC. Access and service use by children with autism spectrum disorders in Medicaid Managed Care. J Autism Dev Disord 2005 Feb;35(1):3-13. [Medline: 15796117]
22. Boisvert M, Lang R, Andrianopoulos M, Boscardin ML. Telepractice in the assessment and treatment of individuals with autism spectrum disorders: A systematic review. Dev Neurorehabil 2010;13(6):423-432. [doi: 10.3109/17518423.2010.499889] [Medline: 20887200]
23. Vismara LA, McCormick C, Young GS, Nadhan A, Monlux K. Preliminary findings of a telehealth approach to parent training in autism. J Autism Dev Disord 2013 Dec;43(12):2953-2969. [doi: 10.1007/s10803-013-1841-8] [Medline: 23677382]
24. Vismara LA, Young GS, Rogers SJ. Telehealth for expanding the reach of early autism training to parents. Autism Res Treat 2012;2012:121878 [FREE Full text] [doi: 10.1155/2012/121878] [Medline: 23227334]
25. Wacker DP, Lee JF, Padilla Dalmau YC, Kopelman TG, Lindgren SD, Kuhle J, et al. Conducting functional communication training via telehealth to reduce the problem behavior of young children with autism. J Dev Phys Disabil 2013 Feb 1;25(1):35-48 [FREE Full text] [doi: 10.1007/s10882-012-9314-0] [Medline: 23543855]
26. Wainer AL, Ingersoll BR. Increasing access to an ASD imitation intervention via a telehealth parent training program. J Autism Dev Disord 2014 Jul 18:1-14. [doi: 10.1007/s10803-014-2186-7] [Medline: 25035089]
27. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, (DSM-5). Arlington, VA: Amer Psychiatric Pub; 2013.
28. Clark BG, Magill-Evans JE, Koning CJ. Youth with Autism Spectrum Disorders: Self- and proxy-reported quality of life and adaptive functioning. Focus Autism Other Dev Disabl 2014 Feb 26;30(1):57-64. [doi: 10.1177/1088357614522289]
29. Allik H, Larsson J, Smedje H. Health-related quality of life in parents of school-age children with Asperger Syndrome or high-functioning autism. Health Qual Life Outcomes 2006;4:1 [FREE Full text] [doi: 10.1186/1477-7525-4-1] [Medline: 16393335]
30. Mugno D, Ruta L, D'Arrigo VG, Mazzone L. Impairment of quality of life in parents of children and adolescents with pervasive developmental disorder. Health Qual Life Outcomes 2007;5:22 [FREE Full text] [doi: 10.1186/1477-7525-5-22] [Medline: 17466072]
31. Maglione MA, Gans D, Das L, Timbie J, Kasari C, Technical EP, HRSA Autism Intervention Research – Behavioral (AIR-B) Network. Nonmedical interventions for children with ASD: Recommended guidelines and further research needs. Pediatrics 2012 Nov;130 Suppl 2:S169-S178 [FREE Full text] [doi: 10.1542/peds.2012-0900O] [Medline: 23118248]
32. National Research Council. Educating children with autism. Washington, DC: National Academy Press; 2001.
33. Rocha ML, Schreibman L, Stahmer AC. Effectiveness of training parents to teach joint attention in children with autism. J Early Interv 2007 Jan 01;29(2):154-172. [doi: 10.1177/105381510702900207]
34. Ingersoll B, Wainer A. Initial efficacy of Project ImPACT: a parent-mediated social communication intervention for young children with ASD. J Autism Dev Disord 2013 Dec;43(12):2943-2952. [doi: 10.1007/s10803-013-1840-9] [Medline: 23689760]
35. Kaiser A, Hancock T, Nietfeld J. Early education and development: The effects of parent-implemented Enhanced Milieu Teaching on the social communication of children who have Autism. Early Educ Dev 2010;11(4):423-446. [doi: 10.1207/s15566935eed1104]
36. Nefdt N, Koegel R, Singer G, Gerber M. The use of a self-directed learning program to provide introductory training in pivotal response treatment to parents of children with autism. J Posit Behav Interv 2009 May 07;12(1):23-32. [doi: 10.1177/1098300709334796]
37. Coolican J, Smith IM, Bryson SE. Brief parent training in pivotal response treatment for preschoolers with autism. J Child Psychol Psychiatry 2010 Dec;51(12):1321-1330. [doi: 10.1111/j.1469-7610.2010.02326.x] [Medline: 21073457]
38. Keen D, Couzens D, Muspratt S, Rodger S. The effects of a parent-focused intervention for children with a recent diagnosis of autism spectrum disorder on parenting stress and competence. Res Autism Spectr Disord 2010 Apr;4(2):229-241. [doi: 10.1016/j.rasd.2009.09.009]
39. Tonge B, Brereton A, Kiomall M, Mackinnon A, King N, Rinehart N. Effects on parental mental health of an education and skills training program for parents of young children with autism: a randomized controlled trial. J Am Acad Child Adolesc Psychiatry 2006 May;45(5):561-569. [doi: 10.1097/01.chi.0000205701.48324.26] [Medline: 16670650]
40. Thomas KC, Ellis AR, McLaurin C, Daniels J, Morrissey JP. Access to care for autism-related services. J Autism Dev Disord 2007 Nov;37(10):1902-1912. [doi: 10.1007/s10803-006-0323-7] [Medline: 17372817]

41. Stahmer AC, Gist K. The effects of an accelerated parent education program on technique mastery and child outcome. J Posit Behav Interv 2001 Apr 01;3(2):75-82. [doi: 10.1177/109830070100300203]
42. Symon JB. Parent education for autism: Issues in providing services at a distance. J Posit Behav Interv 2001 Jul 01;3(3):160-174. [doi: 10.1177/109830070100300304]
43. Taylor TK, Webster-Stratton C, Feil EG, Broadbent B, Widdop CS, Severson HH. Computer-based intervention with coaching: an example using the Incredible Years program. Cogn Behav Ther 2008;37(4):233-246 [FREE Full text] [doi: 10.1080/16506070802364511] [Medline: 18803072]
44. Hartley D. Rural health disparities, population health, and rural culture. Am J Public Health 2004 Oct;94(10):1675-1678. [Medline: 15451729]
45. Hamad CD, Serna RW, Morrison L, Fleming R. Extending the reach of early intervention training for practitioners: A preliminary investigation of an online curriculum for teaching behavioral intervention knowledge in autism to families and service providers. Infants Young Child 2010 Jul 1;23(3):195-208 [FREE Full text] [doi: 10.1097/IYC.0b013e3181e32d5e] [Medline: 23504540]
46. Howroyd C, Peeters T. Parent participation in early intervention with software-assisted guidance from AutismPro. Good Autism Pract 2007;8(2):31-36 [FREE Full text]
47. Jang J, Dixon DR, Tarbox J, Granpeesheh D, Kornack J, de Nocker Y. Randomized trial of an eLearning program for training family members of children with autism in the principles and procedures of applied behavior analysis. Res Autism Spectr Disord 2012 Apr;6(2):852-856. [doi: 10.1016/j.rasd.2011.11.004]
48. Kobak KA, Stone WL, Wallace E, Warren Z, Swanson A, Robson K. A web-based tutorial for parents of young children with autism: results from a pilot study. Telemed J E Health 2011 Dec;17(10):804-808 [FREE Full text] [doi: 10.1089/tmj.2011.0060] [Medline: 22011005]
49. Wainer AL, Ingersoll BR. Disseminating ASD interventions: a pilot study of a distance learning program for parents and professionals. J Autism Dev Disord 2013 Jan;43(1):11-24. [doi: 10.1007/s10803-012-1538-4] [Medline: 22547028]
50. McCullough A. Viability and effectiveness of teletherapy for preschool children with special needs. Int J Lang Commun Disord 2001;36 Suppl:321-326. [Medline: 11340805]
51. Or CKL, Karsh B. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009 Aug;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
52. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
53. Meischke H, Lozano P, Zhou C, Garrison MM, Christakis D. Engagement in "My Child's Asthma": an interactive web-based pediatric asthma management intervention. Int J Med Inform 2011 Nov;80(11):765-774 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.08.002] [Medline: 21958551]
54. Kazdin AE. Assessing the clinical or applied importance of behavior change through social validation. Behav Modif 1977 Oct 01;1(4):427-452. [doi: 10.1177/014544557714001]
55. Barbosa CD, Balp M, Kulich K, Germain N, Rofail D. A literature review to explore the link between treatment satisfaction and adherence, compliance, and persistence. Patient Prefer Adherence 2012;6:39-48 [FREE Full text] [doi: 10.2147/PPA.S24752] [Medline: 22272068]
56. Reimers T, Wacker D, Cooper L, DeRaad A. Clinical evaluation of the variables associated with treatment acceptability and their relation to compliance. Behav Disord 1992;18(1):67-70 [FREE Full text]
57. Kazdin AE. Technology-Based Interventions and Reducing the Burdens of Mental Illness: Perspectives and Comments on the Special Series. Cogn Behav Pract 2015 Aug;22(3):359-366. [doi: 10.1016/j.cbpra.2015.04.004]
58. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems--A meta-analytic review. J Child Psychol Psychiatry 2006 Jan;47(1):99-111. [doi: 10.1111/j.1469-7610.2005.01544.x] [Medline: 16405646]
59. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV TR. Washington, DC: American Psychiatric Association; 2000.
60. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000 Jun;30(3):205-223. [Medline: 11055457]
61. Mullen E. Mullen scales of early learning. Circle Pines, MN: American Guidance Service, Inc; 1995.
62. Ingersoll B, Dvortcsak A. Teaching Social Communication to Children with Autism: A Manual for Parents. NY: The Guilford Press; 2010.
63. Bunz U. The Computer-Email-Web (CEW) Fluency Scale: Development and validation. Int J Hum Comput Interact 2004 Dec;17(4):479-506. [doi: 10.1207/s15327590ijhc1704_3]
64. Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
65. Berger N, Ingersoll B. Psychometric properties and treatment comparisons: Measuring the social validity of skill building interventions for toddlers with autism spectrum disorders. 2014 May Presented at: Poster presented at Autism Speaks' Toddler Treatment Network Pre-Conference, International Meeting for Autism Research; May; 2014; San Sebastian, Spain.

66. Kazdin AE. Acceptability of alternative treatments for deviant child behavior. J Appl Behav Anal 1980;13(2):259-273 [FREE Full text] [doi: 10.1901/jaba.1980.13-259] [Medline: 7380752]
67. Elliott SN, Treuting MVB. The behavior intervention rating scale: Development and validation of a pretreatment acceptability and effectiveness measure. J Sch Psychol 1991 Mar;29(1):43-51. [doi: 10.1016/0022-4405(91)90014-I]
68. Siller M, Morgan L, Turner-Brown L, Baggett KM, Baranek GT, Brian J, et al. Designing studies to evaluate parent-mediated interventions for toddlers with autism spectrum disorder. J Early Interv 2014 Jul 11;35(4):355-377. [doi: 10.1177/1053815114542507]
69. Ingersoll BR, Wainer AL. Pilot study of a school-based parent training program for preschoolers with ASD. Autism 2013 Jul;17(4):434-448. [doi: 10.1177/1362361311427155] [Medline: 22087044]
70. Steele RM, Mummery WK, Dwyer T. Examination of program exposure across intervention delivery modes: face-to-face versus internet. Int J Behav Nutr Phys Act 2007;4:7 [FREE Full text] [doi: 10.1186/1479-5868-4-7] [Medline: 17352817]
71. Carter AS, Messinger DS, Stone WL, Celimli S, Nahmias AS, Yoder P. A randomized controlled trial of Hanen's 'More Than Words' in toddlers with early autism symptoms. J Child Psychol Psychiatry 2011 Jul;52(7):741-752. [doi: 10.1111/j.1469-7610.2011.02395.x] [Medline: 21418212]
72. Kasari C, Gulsrud AC, Wong C, Kwon S, Locke J. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. J Autism Dev Disord 2010 Sep;40(9):1045-1056 [FREE Full text] [doi: 10.1007/s10803-010-0955-5] [Medline: 20145986]
73. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
74. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2 Suppl):TC9-T20. [doi: 10.1097/NCN.0b013e3182155274] [Medline: 21372645]
75. Vismara LA, Colombi C, Rogers SJ. Can one hour per week of therapy lead to lasting changes in young children with autism? Autism 2009 Jan;13(1):93-115. [doi: 10.1177/1362361307098516] [Medline: 19176579]
76. Feil EG, Baggett KM, Davis B, Sheeber L, Landry S, Carta JJ, et al. Expanding the reach of preventive interventions: development of an Internet-based training for parents of infants. Child Maltreat 2008 Nov;13(4):334-346 [FREE Full text] [doi: 10.1177/1077559508322446] [Medline: 18843143]
77. McMahon R, Forehand R, Griest D, Wells K. Who drops out of treatment during parent behavioral training? Behav Couns Q 1981;1(1):79-85.
78. Werba BE, Eyberg SM, Boggs SR, Algina J. Predicting outcome in parent-child interaction therapy: Success and attrition. Behav Modif 2006 Sep;30(5):618-646. [doi: 10.1177/0145445504272977] [Medline: 16894233]
79. Lee H, Friedman ME, Cukor P, Ahern D. Interactive voice response system (IVRS) in health care services. Nurs Outlook 2003;51(6):277-283. [Medline: 14688763]
80. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: A comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
81. Wang J, Senecal S. Measuring perceived website usability. J Internet Commer 2007 Aug 08;6(4):97-112. [doi: 10.1080/15332860802086318]
82. Cranen K, Huis In't Veld R, Ijzerman M, Vollenbroek-Hutten M. Change of patients' perceptions of telemedicine after brief use. Telemed J E Health 2011 Sep;17(7):530-535. [doi: 10.1089/tmj.2010.0208] [Medline: 21767150]
83. Attkisson C, Zwick R. The client satisfaction questionnaire. Eval Program Plann 1982 Jan;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-X]
84. Brestan EV, Jacobs JR, Rayfield AD, Eyberg SM. A consumer satisfaction measure for parent-child treatments and its relation to measures of child behavior change. Behav Ther 1999 Dec;30(1):17-30. [doi: 10.1016/S0005-7894(99)80043-4]
85. Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL, editors. Usability evaluation in industry. London, UK: Taylor & Francis, London; 1996.

Abbreviations

ADOS: Autism Diagnostic Observation Schedule
ANOVA: analysis of variance
ASD: Autism Spectrum Disorder
CBT: Cognitive-Behavioral Therapy
CES-D: Center for Epidemiological Studies-Depression Scale
CEWFS: Computer-Email-Web Fluency Scale
DSM-IV-TR: Diagnostic and Statistical Manual – 4th Edition, Text Revision
MANOVA: multivariate analysis of variance

Edited by P Bamidis; submitted 08.07.15; peer-reviewed by P Yellowlees, E Konstantinidis, A Luneski; comments to author 25.07.15; revised version received 13.08.15; accepted 07.09.15; published 06.10.15.

Please cite as:
Ingersoll B, Berger NI
Parent Engagement With a Telehealth-Based Parent-Mediated Intervention Program for Children With Autism Spectrum Disorders: Predictors of Program Use and Parent Outcomes
J Med Internet Res 2015;17(10):e227
URL: http://www.jmir.org/2015/10/e227/
doi:10.2196/jmir.4913
PMID:26443557

Original Paper

Impact of Educational Level on Study Attrition and Evaluation of Web-Based Computer-Tailored Interventions: Results From Seven Randomized Controlled Trials

Dominique A Reinwand[1,2], MSc; Rik Crutzen[1], PhD; Iman Elfeddali[1,3,4], PhD; Francine Schneider[1], PhD; Daniela Nadine Schulz[1], PhD; Eline Suzanne Smit[1,5], PhD; Nicola Esther Stanczyk[1], PhD; Huibert Tange[6], PhD; Viola Voncken-Brewster[6], MSc; Michel Jean Louis Walthouwer[1], MSc; Ciska Hoving[1], PhD; Hein de Vries[1], PhD

[1]CAPHRI, Department of Health Promotion, Maastricht University, Maastricht, Netherlands

[2]Jacobs Center for Lifelong Learning and Institutional Development, Jacobs University, Bremen, Germany

[3]GGz Breburg, Tilburg, Netherlands

[4]Tranzo Department, Tilburg University, Tilburg, Netherlands

[5]Amsterdam School of Communication Research (ASCoR), Department of Communication Science, University of Amsterdam, Amsterdam, Netherlands

[6]CAPHRI, Department of Family Medicine, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Dominique A Reinwand, MSc
CAPHRI
Department of Health Promotion
Maastricht University
P. O. Box 616
Maastricht, MD 6200
Netherlands
Phone: 31 433882435
Fax: 31 433671032
Email: d.reinwand@maastrichtuniversity.nl

Related Article:

This is a corrected version. See correction statement: http://www.jmir.org/2015/11/e252

Abstract

Background: Web-based computer-tailored interventions have shown to be effective in improving health behavior; however, high dropout attrition is a major issue in these interventions.

Objective: The aim of this study is to assess whether people with a lower educational level drop out from studies more frequently compared to people with a higher educational level and to what extent this depends on evaluation of these interventions.

Methods: Data from 7 randomized controlled trials of Web-based computer-tailored interventions were used to investigate dropout rates among participants with different educational levels. To be able to compare higher and lower educated participants, intervention evaluation was assessed by pooling data from these studies. Logistic regression analysis was used to assess whether intervention evaluation predicted dropout at follow-up measurements.

Results: In 3 studies, we found a higher study dropout attrition rate among participants with a lower educational level, whereas in 2 studies we found that middle educated participants had a higher dropout attrition rate compared to highly educated participants. In 4 studies, no such significant difference was found. Three of 7 studies showed that participants with a lower or middle educational level evaluated the interventions significantly better than highly educated participants ("Alcohol-Everything within the Limit": $F_{2,376}$=5.97, P=.003; "My Healthy Behavior": $F_{2,359}$=5.52, P=.004; "Master Your Breath": $F_{2,317}$=3.17, P=.04). One study found lower intervention evaluation by lower educated participants compared to participants with a middle educational level ("Weight in Balance": $F_{2,37}$=3.17, P=.05). Low evaluation of the interventions was not a significant predictor of dropout at a later follow-up measurement in any of the studies.

Conclusions: Dropout attrition rates were higher among participants with a lower or middle educational level compared with highly educated participants. Although lower educated participants evaluated the interventions better in approximately half of the studies, evaluation did not predict dropout attrition. Further research is needed to find other explanations for high dropout rates among lower educated participants.

(J Med Internet Res 2015;17(10):e228) doi:10.2196/jmir.4941

KEYWORDS

dropout; attrition; educational level; computer tailoring; Web-based intervention; eHealth; evaluation; meta-analysis

Introduction

Previous studies have demonstrated that Web-based computer-tailored interventions can be effective in motivating individuals to adopt different health behaviors [1-3], such as increasing physical activity [4-10], improving healthy nutrition [11-14], smoking cessation [15-18], and reducing alcohol intake [19-21], and they have been successfully applied to multiple health behaviors [22,23]. In addition, these interventions have been found to be more cost-effective than usual care or nontailored information [24-27].

According to Eysenbach [28], dropout, either not completing the study or missing follow-up measurements, is a "fundamental characteristic" of Internet interventions and a problematic issue. The loss of participants to follow-up, dropout attrition, makes analyses and statements of the effectiveness of these interventions more complicated and less valid [28] because most outcome measures are assessed during follow-up and dropout from the intervention seems to be related to dropout attrition [29]. Therefore, it is important to find out why participants do not complete Web-based studies to ultimately understand and reduce this problem.

Acquiring follow-up measurements from at-risk groups, such as people with a lower educational level, is important because unhealthy lifestyle behaviors are most common among people with a lower educational level [30-32]. Lower educated people are known to eat less fruits and vegetables [33,34], are less physically active [30,35], consume alcohol more often in unhealthy quantities [36], use more tobacco [37-39], and have a higher likelihood of being obese [40] compared to highly educated people. It is not only important to reach this group with Web-based computer-tailored interventions, but also to prevent lower educated participants from dropping out of the follow-up measurements to be able to collect information about the effectiveness of the intervention [29].

Christensen and Mackinnon [41] already raised the issue of insufficient research regarding study dropout in 2006. Since then, findings about dropout among participants with different educational levels are still rarely reported and show ambiguous results. Although some studies revealed that people with a lower educational level have higher dropout rates in Web-based computer-tailored interventions than highly educated people [42-46], other studies did not find educational differences in terms of dropout [46-49]. To the best of our knowledge, no study indicates a significantly higher dropout among highly educated participants, but there remains a need to obtain more insight into dropout among people with different educational levels to be able to reduce dropout.

The reason for dropout among lower educated people is discussed rarely in the literature. Possible reasons for dropout can be intervention characteristics (eg, workload, content), personal characteristics such as educational level, or it can be related to participants' perceptions of the interventions, such as a lack of perceived benefit, which may result in dissatisfaction [50]. Dissatisfaction with the intervention can be reflected in the evaluation of the intervention. It has been shown that participants who do evaluate the intervention as less positive are more likely to drop out [51] and, therefore, might not be interested in attending follow-up measurements. In other words, evaluation might be a predictor of dropout attrition in Web-based computer-tailored interventions, but little thought has been given to this aspect, which makes it difficult to draw valid conclusions [52].

Therefore, the aim of this study is first to examine if the dropout attrition rates in our 7 randomized controlled trials (RCTs) of Web-based computer-tailored interventions were higher for people with a lower educational level than people with a middle or high educational level. Second, we assessed whether people with different educational levels evaluated these interventions differently. Finally, we analyzed whether participations' evaluation of the interventions predicted dropout at subsequent follow-up measurements.

Methods

Studies

To examine differences in dropout attrition and evaluation between participants with different educational levels, we used a convenience sample of participants from 7 Web-based computer-tailored intervention studies that were conducted in the past years (2010-2014) at the Department of Health Promotion of Maastricht University in the Netherlands.

The studies were RCTs to evaluate interventions that used computer-tailored techniques to improve diverse health behaviors. The study "Master Your Breath" (MYB) focused on increasing physical activity and smoking cessation among chronic obstructive pulmonary disease (COPD) patients and people at risk for COPD. The 3 studies "Stay Quit for You" (SQ4U), "Support to Quit" (STQ), and "Personal Advice in Stopping smoking" (PAS) focused on smoking cessation. "Weight in Balance" (WIB) aimed to prevent obesity by targeting physical activity and energy intake. The study "My Healthy Behavior" (MHB) targeted the following health behaviors: physical activity, fruit and vegetable consumption,

alcohol intake, and smoking. The study "Alcohol-Everything within the Limits" (AEL) focused on moderate alcohol intake and is the only study that was not carried out in the Netherlands but in Germany.

All selected studies made use of the I-Change model [53,54], which postulates that the behavior change process has at least 3 phases: awareness, motivation, and action. The first factor is determined by factors such as behavioral awareness, knowledge, and risk perceptions. The second phase is determined by attitudes, social influence beliefs, and self-efficacy expectations, and results in a certain intention to perform a particular behavior. The third factor is determined by self-efficacy, action planning, skills, and barriers. The tailored feedback messages of the studies included in this paper have a strong focus on inter alia these determinants. A detailed description of these RCTs and the related publications can be found in Table 1.

Table 1. Summary of the Web-based computer-tailored interventions.

Study	Reference	Participants	Study groups	Intervention	Follow-up
AEL	Design and effects: [20]	German general population aged 18-69 years	Two intervention groups that differed in the computer-tailored feedback strategies (alternating vs summative) compared to 1 control group that received no computer-tailored feedback.	A 3-session, Web-based computer-tailored intervention aiming to reduce alcohol intake in high-risk adult drinkers.	T1=3 months; T2=6 months
MHB	Study protocol: [55]; effects: [20,26]	Dutch general population aged 19-65 years	Two experimental groups (ie, a sequential behavior tailoring condition and a simultaneous behavior tailoring condition) and 1 control group that that received only a tailored health risk appraisal but no motivational computer-tailored feedback.	Five lifestyle behaviors of smoking, alcohol intake, fruit consumption, vegetable consumption, and physical activity addressing computer-tailored feedback at several times.	T1=12 months; T2=24 months
MYB	Study protocol: [56]; effects: [57]	People with or at risk for COPD in the Netherlands	One intervention group received Web-based computer-tailored self-management intervention; the control group received usual care.	Web-based, computer-tailored self-management intervention with the aim to increase physical activity and support smoking cessation.	T1=6 months
PAS	Study protocol: [58]; effects: [18,25,59]	Adult Dutch smokers with intention to stop smoking within 6 months	Intervention group with computer-tailored information to quit smoking compared to control group that received no computer-tailored feedback.	A Web-based computer-tailored smoking cessation intervention.	T1=6 weeks; T2=6 months; T3=12 months
STQ	Study protocol: [60]; effects: [61,62]	Dutch smokers who were motivated to stop smoking and aged ≥18 years	Intervention groups 2 (video/text) × (low/middle/high socioeconomic status). Respondents were assigned to 1 of the intervention groups (text- vs video-tailored feedback) or to the control group (nontailored generic advice).	Comparing Web-based text and a Web-based video-driven computer-tailored approach for low and high SES smokers, this incorporates multiple computer-tailored feedback moments with the aim to support smoking cessation.	T1=6 months; T2=12 months
SQ4U	Study protocol: [63]; effects: [15]	Dutch daily smokers aged 18-65 years who were motivated to stop smoking	Two intervention groups (Action Plan, Action Plan+), 1 control group that received no computer-tailored feedback.	Two computer-tailored interventions to prevent smoking relapse. Provides tailored feedback in the Action Plan+ group after stop smoking attempts, in the Action Plan group after T0 measurement.	T1=6 months; T2=12 months
WIB	Study protocol: [64]; effects: [65]	Normal and overweight adults from the Netherlands	Two intervention groups (video and text) and 1 waiting list control group.	Computer-tailored feedback via text or video to prevent weight gain or support modest weight loss by targeting physical activity and energy intake.	T1=6 months

Measurement

In all 7 studies, educational level was assessed by asking participants about their highest completed level of education. In-line with national guidelines, educational level was categorized into 3 groups: lower (1=no education, primary or lower vocational school), middle (2=secondary vocational school or high school), and higher (3=higher professional education or university) educational level [66].

All studies included a process evaluation assessment to evaluate the intervention among participants within the intervention group. Participants were asked to evaluate the tailored feedback and the intervention. The process evaluation assessments of the 7 studies included had one item in common that asked participants to grade the intervention that they participated in: "Please evaluate the intervention with a school grade from 1 to 10" (10=highest grade, 1=lowest grade according to the Dutch school grading system; AEL: 15=highest grade, 1=lowest grade, which is in-line with the German school grading system).

To assess dropout attrition, participants who completed the baseline measurement but did not complete the follow-up measurement were characterized as dropouts (1=dropout; 0=completed follow-up). We assessed dropout attrition within differently educated participants for each follow-up measurement separately. Furthermore, we used the last available evaluation moment as predictor of dropout for the following measurement. Table 1 gives an overview of the specific follow-up moments per study.

Statistical Analysis

All analyses were performed with SPSS 20.0 (IBM Corp, Armonk, NY, USA). Descriptive statistics were used to describe sample characteristics. Per study, a logistic regression analysis was conducted to examine if dropout rates differed for each educational level. To control for multiple testing, the Benjamini and Hochberg linear step-up method was used for each study [67,68]. With the use of an Excel template, the adjusted significance levels were calculated [69].

Differences between the educational levels with regard to evaluation of the Web-based computer-tailored interventions were analyzed by means of ANOVAs and Tukey honestly significant difference (HSD) tests. Control groups were excluded from analysis with regard to evaluation of the intervention because they could not evaluate it.

To be able to give a more general picture of whether lower and higher educated participants from the intervention groups evaluated Web-based computer-tailored interventions differently, the Exploratory Software for Confidence Intervals (ESCI) Excel template [70] was used for pooling the data by means of a meta-analysis (Table 1). The meta-analysis used a random effect model and gave an impression of the overall differences for intervention evaluation between lower and higher educated participants (ie, by subtracting the evaluation of the most different groups, the lower educated participants from the higher educated participants). In one study (MHB), the evaluation item was assessed at multiple follow-up measurements; in this case, we included only the last follow-up measurement [71].

Finally, logistic regression analyses were conducted to examine if dropout was predicted by evaluation in 4 of the 7 interventions among participants with different educational levels. We excluded the studies WIB, PAS, and MYB from this analysis because their evaluation assessment took place during the last follow-up measurement; therefore, it was not possible to assess evaluation as a predictor of dropout in these studies. To identify possible interaction effects, an interaction term of educational level and evaluation was used in the regression model. If this interaction term was significant, then analyses were conducted separately per educational level. Analyses were corrected for age and gender. A *P* value of .05 was used as the significance level for all analysis.

Results

Participants

Table 2 shows the educational level, mean age, and gender distribution of the participants within the 7 selected studies at baseline.

Table 2. Baseline sample characteristics of the participants in the Web-based computer-tailored interventions.

Study	N	Educational level, n (%)[a]			Age (years), mean (SD)	Gender (male), n (%)
		Low	Middle	High		
AEL	1149	483 (44.8)	256 (23.8)	338 (31.4)	43.82 (15.51)	550 (47.9)
MHB	5055	515 (10.4)	2334 (47.1)	2112 (42.6)	44.15 (12.67)	2661 (52.6)
MYB	1307	386 (29.5)	427 (32.7)	494 (37.8)	57.64 (7.22)	627 (47.9)
PAS	1123	238 (21.2)	513 (45.7)	372 (33.1)	49.47 (32.55)	535 (47.6)
STQ	2099	707 (33.6)	782 (37.3)	612 (29.2)	45.33 (13.21)	821 (39.1)
SQ4U	2031	207 (10.2)	1130 (55.6)	694 (34.2)	40.88 (11.80)	766 (37.7)
WIB	1419	214 (15.1)	436 (30.7)	769 (54.2)	48.13 (11.52)	588 (41.4)

[a] For reference, the average educational level in Germany for low, middle, and high is 39, 22, and 27, respectively [72]; for the Netherlands, it is 30, 28, and 42, respectively [73].

Dropout

Table 3 shows the results of the dropout analyses with regard to the educational level for each study, each follow-up measurement including the dropout rates, and study group in detail with high education as the reference group.

Table 3. Results of a logistic regression examine dropout attrition among different educational groups.

Study, follow-up, and group	Dropout, n (%)	Educational level[a]			
		Low		Middle	
		OR (95% CI)	*P*	OR (95% CI)	*P*
AEL					
T1	398 (34.6)				
Sequential		0.61 (0.10-3.59)	.58	1.03 (0.15-7.18)	.97
Simultaneously		1.65 (0.61-4.58)	.32	—[c]	.99
Control		—[c]	.99	—[c]	.99
T2	436 (37.9)				
Sequential		1.16 (0.24-5.52)	.89	1.90 (0.18-19.37)	.58
Simultaneously		1.15 (0.34-3.84)	.81	1.23 (0.21-7.13)	.81
Control		0.90 (0.11-7.06)	.92	0.51 (0.37-7.09)	.61
MHB					
T1	3317 (65.6)				
Sequential		1.52 (1.14-2.01)	.004[b]	1.05 (0.83-1.32)	.68
Simultaneously		1.57 (1.18-2.08)	.002[b]	1.39 (1.09-1.78)	.007[b]
Control		1.32 (1.00-1.73)	.04	1.27 (1.00-1.61)	.04
T2	3602 (71.3)				
Sequential		1.43 (1.06-1.94)	.01[b]	0.95 (0.74-1.23)	.73
Simultaneously		1.51 (1.12-2.04)	.006[b]	1.50 (1.16-1.94)	.002[b]
Control		1.23 (0.93-1.62)	.14	0.97 (0.76-1.23)	.81
MYB					
T1	254 (19.4)				
Intervention		1.33 (0.84-2.12)	.21	1.40 (0.89-2.20)	.13
Control		1.14 (0.67-1.95)	.61	1.17 (0.70-1.96)	.53
PAS					
T1	674 (60.0)				
Control		0.93 (0.58-1.49)	.77	0.88 (0.59-1.31)	.55
Tailoring only		2.02 (1.23-3.33)	.005	1.37 (0.93-2.01)	.10
T2	831 (74.0)				
Control		0.93 (0.54-1.56)	.77	0.89 (0.57-1.39)	.61
Tailoring only		2.04 (1.15-3.60)	.01	1.41 (0.93-2.15)	.10
T3	967 (86.1)				
Control		1.42 (0.71-2.86)	.32	0.89 (0.52-1.52)	.67
Tailoring only		1.41 (0.67-2.97)	.35	1.03 (0.60-1.78)	.90
STQ					
T1	1306 (62.2)				
Video		1.90 (1.24-2.90)	.003[b]	1.39 (0.93-2.09)	10
Text		1.29 (0.87-1.91)	.19	1.22 (0.83-1.78)	.29
Control		0.98 (0.67-1.47)	.98	0.71 (0.49-1.05)	.09
T2	1437 (68.5)				

Study, follow-up, and group	Dropout, n (%)	Educational level[a]			
		Low		Middle	
		OR (95% CI)	*P*	OR (95% CI)	*P*
Video		1.95 (1.26-3.02)	.003[b]	1.39 (0.92-2.09)	.11
Text		2.31 (1.52-3.51)	.<001[b]	1.29 (0.88-1.89)	.18
Control		1.36 (0.90-2.04)	.13	1.24 (0.84-1.84)	.66
SQ4U					
T1	1251 (61.9)				
Action Plan		1.26 (0.75-2.12)	.36	1.03 (0.74-1.44)	.83
Action Plan +		1.71 (0.92-3.18)	.08	1.14 (0.81-1.60)	.44
Control		1.73 (0.91-3.27)	.09	0.90 (0.63-1.28)	.57
T2	1465 (72.1)				
Action Plan		2.33 (1.24-4.35)	.01	1.30 (0.91-1.86)	.14
Action Plan +		2.25 (1-11-4.52)	.02	1.55 (1.07-2.24)	.01
Control		2.00 (1.02-3.92)	.04	1.35 (0.94-1.94)	.09
WIB					
T1	404 (28.5)				
Video		1.50 (0.87-2.59)	.15	1.23 (0.81-2.01)	.29
Text		2.29 (1.33-3.95)	.003[b]	1.12 (0.74-1.69)	.60
Control		1.57 (0.81-3.04)	.18	2.01 (1.22-3.32)	.006[b]

[a] All analysis are corrected for age and gender. High education is the reference group.

[b] Significant *P* values after correction for multiple comparisons according to Benjamini-Hochberg.

[c] Odds ratios are not reported due to low cell count.

After correction for multiple testing, significantly higher dropout rates were found within 3 studies (MHB, STQ, WIB) among lower educated participants compared to higher educated ones. Furthermore, in these 3 studies, dropout attrition was also significantly higher among middle educated participants in comparison with higher educated participants. In 4 of 7 studies (AEL, MYB, PAS, SQ4U), no difference in dropout with regard to educational level was found.

Evaluation

Table 4 presents differences between the educational groups with regard to evaluation of the Web-based computer-tailored interventions in detail. In 3 of 7 studies (AEL, MHB, MYB), lower educated participants evaluated the intervention significantly higher compared to their counterparts. In one study (WIB), lower educated participants evaluated the intervention less positively compared to middle educated participants.

Table 4. Evaluation of the 7 Web-based computer-tailored interventions by different educational levels.

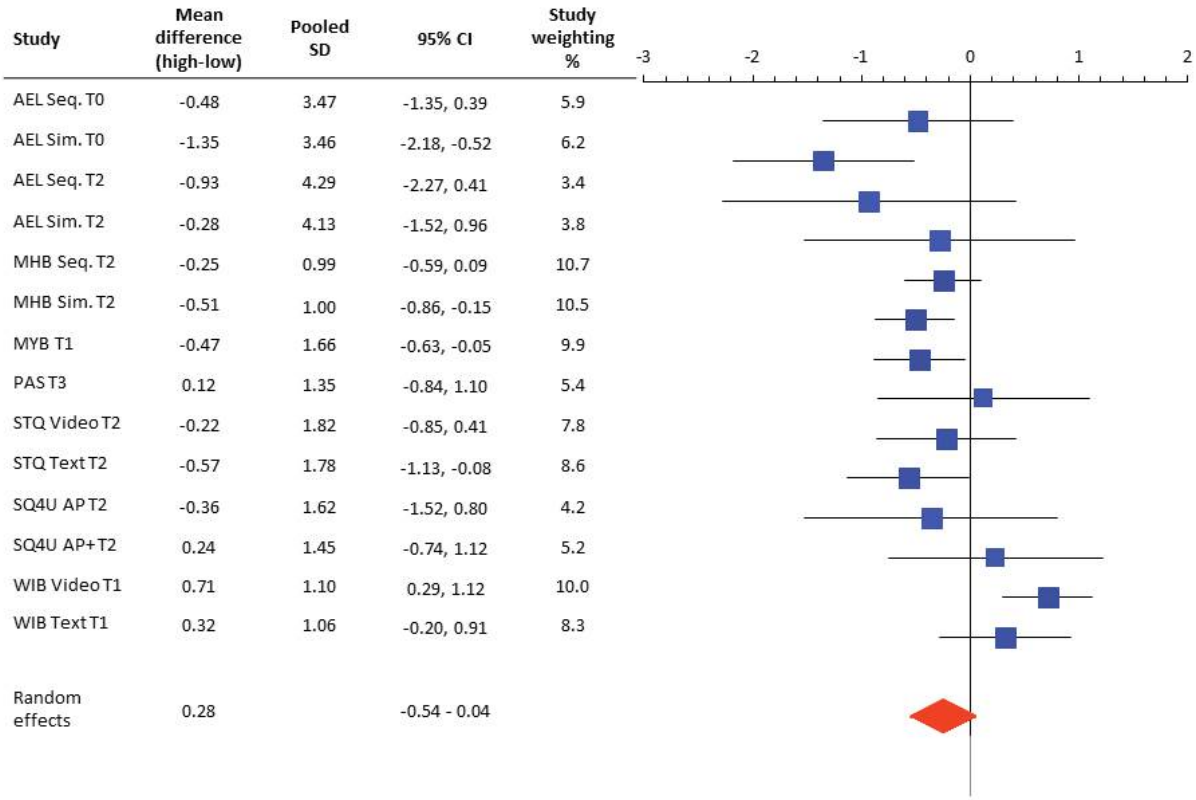
Study and group[a]	Level of education, mean (SD)			*F* (*df1*,*df2*)	*P*	Tukey HSD, *P*		
	Low	Middle	High			L-M[b]	L-H	M-H
AEL								
T0								
Sequential	11.20 (3.48)	11.12 (3.00)	10.72 (3.62)	0.71 (2,340)	.56	.97	.47	.70
Simultaneously	11.75 (3.37)	11.49 (2.85)	10.40 (3.61)	5.97 (2,376)	.003	.82	.002	.05
T2								
Sequential	11.32 (4.27)	11.27 (3.69)	10.39 (4.33)	0.58 (2,229)	.56	.77	.89	.53
Simultaneously	11.09 (4.28)	11.53 (3.30)	10.81 (3.89)	1.26 (2,246)	.56	.99	.29	.48
MHB								
T1								
Sequential	7.43 (1.08)	7.14 (1.79)	7.03 (1.13)	2.12 (2,201)	.12	.98	.48	.11
Simultaneously	7.80 (0.91)	6.94 (1.70)	6.56 (2.41)	1.30 (2,178)	.27	.60	.27	.62
T2								
Sequential	7.78 (1.27)	7.59 (0.94)	7.53 (0.91)	1.04 (2,367)	.35	.64	.37	.71
Simultaneously	7.94 (0.89)	7.68 (0.92)	7.43 (1.02)	5.52 (2,359)	.004	.30	.01	.05
MYB								
T1								
Intervention	7.07 (1.50)	6.93 (1.23)	6.60 (1.77)	3.17 (2,317)	.04	.77	.05	.17
PAS								
T3								
Tailoring only	6.09 (1.70)	6.72 (1.25)	7.03 (1.20)	2.42 (2,81)	.09	.92	.96	.63
STQ								
T2								
Video	6.48 (1.97)	6.18 (2.17)	6.28 (1.64)	0.38 (2,193)	.69	.67	.83	.95
Text	6.53 (1.84)	6.51 (1.23)	5.96 (1.72)	3.29 (2,234)	.04	.99	.08	.07
SQ4U								
T1								
Action Plan	6.56 (1.74)	6.63 (1.41)	6.20 (1.60)	1.29 (2,134)	.28	.98	.79	.25
Action Plan+	6.27 (2.10)	6.49 (1.69)	6.51 (1.27)	0.10 (2,108)	.90	.97	.85	.99
WIB								
T1								
Video	6.99 (1.23)	7.56 (0.78)	7.36 (1.08)	3.17 (2,37)	.05	.05	.27	.21
Text	6.79 (0.92)	7.32 (0.82)	7.11 (1.24)	0.66 (2,51)	.52	.49	.67	.88

[a] T specifies the time of the evaluation measurement.

[b] Level of education: L=low, M=middle, H=high.

Furthermore, the meta-analysis of the 7 studies comparing evaluation of lower and higher educated participants indicated that participants with a lower educational level evaluated the interventions significantly more positively compared to highly educated participants (OR 0.28, 95% CI –0.54 to 0.04, $P<.001$) (see Figure 1). Nevertheless, the meta-analysis revealed presence of a moderate level of heterogeneity (I^2=66.10%) [74,75], which indicates variation across the studies.

Figure 1. Forest plot of mean differences by random effect model of evaluation of Web-based computer-tailored interventions between highly and lower educated participants. Random effects represent the combined effect.



Association of Education and Evaluation with Dropout Attrition

For the 4 studies that evaluated the intervention before the follow-up measurements, a significant interaction between education and evaluation regarding dropout at follow-up was not found (Table 5). Only within the MHB study was a positive association found between the intervention evaluation and educational level. Participants with a middle educational level were more likely to dropout than participants with a higher educational level.

Table 5. Association of education and evaluation with dropout attrition at follow-up.

Study and variables[a]	β	*P*	OR (95% CI)	χ^2_7	R^2
AEL T1				9.0	.054
Education low	–1.02	.35	0.35 (0.04-3.14)		
Education middle	–0.08	.96	0.92 (0.02-37.18)		
Evaluation T0	0.03	.69	1.03 (0.88-1.19)		
Education × evaluation		.48			
MHB T1				96.9	.053
Education low	1.90	.05	6.71 (0.97-46.43)		
Education middle	0.93	.07	2.55 (0.89-7.27)		
Evaluation T0	–0.01	.84	0.99 (0.90-1.08)		
Education × evaluation		.18			
STQ T2				16.6	.033
Education low	0.62	.29	1.86 (0.57-6.00)		
Education middle	0.28	.64	1.32 (0.40-4.35)		
Evaluation T1	–0.49	.52	0.95 (0.81-1.10)		
Education × evaluation		.99			
SQ4U T2				5.6	.051
Education low	4.33	.11	76.36 (0.34-16.713.27)		
Education middle	1.32	.55	3.75 (0.46-303.55)		
Evaluation T1	0.11	.69	1.11 (0.64-1.92)		
Education × evaluation		.40			

[a] T indicates follow-up; high education is reference group. All analyses are corrected for age and gender.

Discussion

Dropout Attrition

The first aim of this study was to evaluate whether participants with a lower educational level have higher dropout attrition from Web-based computer-tailored studies than participants with a medium or high educational level. In 3 of these studies, lower and middle educated participants dropped out more frequently compared to higher educated participants.

A possible explanation for the higher dropout rates may be that lower educated participants tend to use written health information more often [76] and spend less time online seeking health information [77,78]. It could be possible that they lose interest in the intervention sooner, which causes them to drop out of the study.

Also, the fact that lower educated people have an unhealthier lifestyle [30-32] might play a role in dropout. Due to tailoring, participants with an unhealthier lifestyle in multiple health behavior interventions received more recommendations to change their health behavior(s) and this has been found to decrease participants' motivation to change [79]. It might be possible that lower educated participants started the intervention with the aim to change their health behavior, but that receiving information about extensive required changes decreased their self-efficacy to be able to change [80] and could subsequently have decreased their motivation to participate. Another explanation could be that lower educated participants might have been less likely to change their behavior and, thus, may have perceived the recommendations as less feasible, which caused them to drop out of the study [28,81,82]. This could have caused not only usage and nonusage attrition, but also dropout attrition because these 2 kinds of attrition seem to be related [46].

Moreover, lower educated people might be less familiar with Web-based computer-tailored interventions [83,84] and that might result in lower confidence in the effectiveness of those interventions (ie, lower perceived efficacy) and, in turn, could cause an increase in dropout [85,86]. Although these are reasons for nonusage attrition (not using the intervention), it seems convincing that this correlates with dropout attrition because participants who did not evaluate the intervention positively might have little interest in completing follow-up measurements [51].

All participants were asked to complete long questionnaires and received tailored feedback, which must be cognitively processed and requires intensive cognitive performance. Lower educated adults have been shown to have a lower level of health literacy [84,87,88]. They have more difficulties processing new information and this could cause ego depletion [89,90]. Ego depletion may, in turn, reduce the willingness to participate any longer within the study.

Some studies found that dropout attrition could be increased by sending reminders and prompts [78,91-93]. Further research is necessary to evaluate if this is also effective for people with different educational levels.

Evaluation

Against our initial expectation that lower educated participants might evaluate the interventions less positively, we found that lower educated participants evaluated the intervention in 3 of 7 studies more positively compared to their higher educated counterparts. This might be explained by the finding that highly educated people make more use of the Internet as health information resource, whereas these interventions might be newer and more interesting for lower educated people. A review supported this assumption because it shows that people with a high educational level may make more intensive use of several sources (eg, people form their social network, mass media, health professionals) to gain health-related information compared to lower educated people and they might read the received information more superficially [94]. This might result in less elaboration of the messages and a lower evaluation regarding the novelty of the messages. Due to the use of several resources, highly educated participants also rely less on online information and have lower levels of trust in them, which may negatively influence their evaluation of the intervention [95].

Evaluation was not a significant predictor of dropout at follow-up in any of our studies. This suggests that other factors must be important in explaining why participants did not return to the study for follow-up questionnaires. Dropout analysis performed within these studies has shown that a lower educational level, unhealthy lifestyle, low intention to change the behavior, and low self-efficacy were predictors of dropout [20,96,97].

Limitations and Strengths

First, the only item all studies had in common concerning the evaluation of the interventions was an overall grade participants assigned to the intervention. Although we can assume that this item gives an overall impression of participants' evaluation, it might be that participants with different educational levels liked and disliked different aspects of the intervention (eg, layout, provided information, or personal relevance), which was not reflected in this overall grade. However, the evaluation of these different aspects was not equally assessed in all 7 studies. Furthermore, it is possible that participants who did not like the intervention dropped out before completing the evaluation item. In this study, we included only those participants that completed the evaluation item and assessed follow-up at the subsequent measurement.

Second, all interventions were based on the I-Change model and targeted the same social cognitive determinants to change behavior, which allows for comparing the 7 studies. However, a generalization of the results for other interventions must be done with caution because some interventions also used other theories, such as self-regulation theories, as a framework for the educational content.

Although the restricted number of studies is a limitation, including other Web-based tailored interventions might have resulted in even higher program heterogeneity and would have made comparisons even more difficult and results (partly) dependent on program characteristics. Also Wienert and Kuhlmann [98] have determined that tailored interventions are difficult to compare, whereas the interventions included in this study were comparable because they all were based on the same theoretical background (the I-Change model), all 7 studies provided tailored feedback on social cognitive determinants from this model and provided feedback, and all 7 studies included the same program evaluation item. The comparison of interventions using other tailoring techniques than the interventions described in this study is difficult and access to the original data at the individual level would be necessary for further research and adequate analysis [99].

One of the strengths of this study is the access to 7 datasets (at participant level), which allowed us to conduct the analysis with the original data. Second, all studies used at least one identical item to assess the study evaluation which enables us to compare these studies. Finally, all studies had a large number of participants ranging from 1149 to 5055, which makes our results meaningful.

Conclusion

This study showed that for 3 of 7 studies on computer-tailored interventions, participants with a lower educational level dropped out more often from follow-up measurements and tended to evaluate the interventions better compared to participants with a middle and higher educational level. However, the evaluation of the intervention did not predict participation or nonparticipation at follow-up. Based on our results, it is hard to say what other factors may play a role in dropout attrition from Web-based computer-tailored interventions. Further studies might evaluate different aspects of the intervention, besides only the participants' grades, to find more relevant aspects of intervention evaluation.

Future studies should take high dropout among lower educated participants into consideration when developing strategies to decrease high dropout from Web-based computer-tailored interventions.

Acknowledgments

We would like to thank all doctoral advisors who supported the development of the mentioned Web-based computer-tailored interventions: Catherine Bolman, Stef Kremers, Lilian Lechner, Anke Oenema, Liesbeth van Osch, Jean Muris, and Trudy van der Weijden.

Authors' Contributions

All authors contributed to this paper. DR conducted the analyses and drafted the paper. HdV and RC contributed to the design of the paper. All authors contributed to the interpretation of the data and to the writing of the paper. All authors revised the manuscript critically for important intellectual content and read and approved the final manuscript.

Conflicts of Interest

HdV is the scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools. The other authors declare that they have no competing interests.

References

1. Noar SM, Harrington NG, editors. eHealth Applications: Promising Strategies for Behavior Change (Routledge Communication Series). New York: Routledge; 2012.
2. Noar S, Grant Harrington N, Van Stee SK, Shemanski Aldrich R. Tailored health communication to change lifestyle behaviors. American Journal of Lifestyle Medicine 2010 Nov 19;5(2):112-122. [doi: 10.1177/1559827610387255]
3. Wantland D, Portillo C, Holzemer W, Slaughter R, McGhee E. 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]
4. Ammann R, Vandelanotte C, de Vries H, Mummery W. Can a website-delivered computer-tailored physical activity intervention be acceptable, usable, and effective for older people? Health Educ Behav 2013 Apr;40(2):160-170. [doi: 10.1177/1090198112461791] [Medline: 23077157]
5. van 't Riet J, Ruiter RA, Werrij M, de Vries H. Investigating message-framing effects in the context of a tailored intervention promoting physical activity. Health Educ Res 2010 Apr;25(2):343-354 [FREE Full text] [doi: 10.1093/her/cyp061] [Medline: 19841041]
6. Zuercher A. A look at the latest AIDS projections for the United States. Health Aff (Millwood) 1990;9(2):163-170 [FREE Full text] [Medline: 2163963]
7. Norman G, Zabinski M, Adams M, Rosenberg D, Yaroch A, Atienza A. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
8. 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]
9. Peels D, van Stralen MM, Bolman C, Golsteijn RH, de Vries H, Mudde A, et al. The differentiated effectiveness of a printed versus a Web-based tailored physical activity intervention among adults aged over 50. Health Educ Res 2014 Oct;29(5):870-882 [FREE Full text] [doi: 10.1093/her/cyu039] [Medline: 24980023]
10. Konstantinidis E, Billis A, Mouzakidis C, Zilidou V, Antoniou P, Bamidis P. Design, implementation and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. IEEE J. Biomed. Health Inform 2014 Dec 09;PP(99):1-1 [FREE Full text] [doi: 10.1109/JBHI.2014.2378814]
11. Brug J, Campbell M, van Assema P. The application and impact of computer-generated personalized nutrition education: a review of the literature. Patient Educ Couns 1999 Feb;36(2):145-156. [Medline: 10223019]
12. Krebs P, Prochaska J, Rossi J. 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]
13. Hamel L, Robbins L. Computer- and web-based interventions to promote healthy eating among children and adolescents: a systematic review. J Adv Nurs 2013 Jan;69(1):16-30. [doi: 10.1111/j.1365-2648.2012.06086.x] [Medline: 22757605]
14. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
15. Elfeddali I, Bolman C, Candel Math JJ, Wiers R, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]
16. Smit E, de Vries H, Hoving C. Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers. J Med Internet Res 2012;14(3):e82 [FREE Full text] [doi: 10.2196/jmir.1812] [Medline: 22687887]
17. McClure J, Shortreed S, Bogart A, Derry H, Riggs K, St John J, et al. The effect of program design on engagement with an internet-based smoking intervention: randomized factorial trial. J Med Internet Res 2013;15(3):e69 [FREE Full text] [doi: 10.2196/jmir.2508] [Medline: 23529377]
18. Smit E, de Vries H, Hoving C. Effectiveness of a Web-based multiple tailored smoking cessation program: a randomized controlled trial among Dutch adult smokers. J Med Internet Res 2012;14(3):e82 [FREE Full text] [doi: 10.2196/jmir.1812] [Medline: 22687887]

19. van der Wulp NY, Hoving C, Eijmael K, Candel Math JJ, 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]
20. Schulz D, Candel M, Kremers S, Reinwand D, Jander A, de Vries H. Effects of a Web-based tailored intervention to reduce alcohol consumption in adults: randomized controlled trial. J Med Internet Res 2013;15(9):e206 [FREE Full text] [doi: 10.2196/jmir.2568] [Medline: 24045005]
21. Boon B, Risselada A, Huiberts A, Riper H, Smit F. Curbing alcohol use in male adults through computer generated personalized advice: randomized controlled trial. J Med Internet Res 2011;13(2):e43 [FREE Full text] [doi: 10.2196/jmir.1695] [Medline: 21719412]
22. van Keulen HM, Mesters I, Ausems M, van Breukelen G, Campbell M, Resnicow K, et al. Tailored print communication and telephone motivational interviewing are equally successful in improving multiple lifestyle behaviors in a randomized controlled trial. Ann Behav Med 2011 Feb;41(1):104-118 [FREE Full text] [doi: 10.1007/s12160-010-9231-3] [Medline: 20878293]
23. 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]
24. van Keulen HM, Bosmans J, van Tulder MW, Severens J, de Vries H, Brug J, et al. Cost-effectiveness of tailored print communication, telephone motivational interviewing, and a combination of the two: results of an economic evaluation alongside the Vitalum randomized controlled trial. Int J Behav Nutr Phys Act 2010;7:64 [FREE Full text] [doi: 10.1186/1479-5868-7-64] [Medline: 20815869]
25. Smit E, 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]
26. Schulz D, Smit E, Stanczyk N, Kremers SP, de Vries VH, 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]
27. Stanczyk N, Bolman C, Smit E, Candel MJ, Muris JW, de Vries H. How to encourage smokers to participate in web-based computer-tailored smoking cessation programs: a comparison of different recruitment strategies. Health Educ Res 2014 Feb;29(1):23-40 [FREE Full text] [doi: 10.1093/her/cyt104] [Medline: 24287402]
28. 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]
29. Cunningham A. Intervention adherence is related to participant retention: implications for research. J Med Internet Res 2014;16(5):e133 [FREE Full text] [doi: 10.2196/jmir.2948] [Medline: 24849941]
30. Vlismas K, Stavrinos V, Panagiotakos D. Socio-economic status, dietary habits and health-related outcomes in various parts of the world: a review. Cent Eur J Public Health 2009 Jun;17(2):55-63. [Medline: 19662821]
31. Margolis R. Educational differences in healthy behavior changes and adherence among middle-aged Americans. J Health Soc Behav 2013;54(3):353-368 [FREE Full text] [doi: 10.1177/0022146513489312] [Medline: 23988727]
32. Cowell A. The relationship between education and health behavior: some empirical evidence. Health Econ 2006 Feb;15(2):125-146. [doi: 10.1002/hec.1019] [Medline: 16167321]
33. De Irala-Estévez J, Groth M, Johansson L, Oltersdorf U, Prättälä R, Martínez-González MA. A systematic review of socio-economic differences in food habits in Europe: consumption of fruit and vegetables. Eur J Clin Nutr 2000 Sep;54(9):706-714. [Medline: 11002383]
34. Hulshof KF, Brussaard J, Kruizinga A, Telman J, Löwik MR. Socio-economic status, dietary intake and 10 y trends: the Dutch National Food Consumption Survey. Eur J Clin Nutr 2003 Jan;57(1):128-137. [doi: 10.1038/sj.ejcn.1601503] [Medline: 12548307]
35. Gidlow C, Johnston L, Crone D, Ellis N, James D. A systematic review of the relationship between socio-economic position and physical activity. Health Education Journal 2006 Dec 01;65(4):338-367. [doi: 10.1177/0017896906069378]
36. Groot W, Maassen van den Brink M. What does education do to our health? In: Measuring the Effects of Education on Health and Civic Engagement: Proceedings of the Copenhagen Symposium. 2006 Presented at: Copenhagen Symposium on measuring the effects of education on health and civic engagement; Mar 2006; Copenhagen p. 355-364 URL: http://www.oecd.org/edu/innovation-education/37437718.pdf
37. Hiscock R, Bauld L, Amos A, Fidler J, Munafò M. Socioeconomic status and smoking: a review. Ann N Y Acad Sci 2012 Feb;1248:107-123. [doi: 10.1111/j.1749-6632.2011.06202.x] [Medline: 22092035]
38. Savage C. Alcohol and tobacco related health inequity: a population health perspective. J Addict Nurs 2012 Feb;23(1):72-74. [doi: 10.3109/10884602.2011.648739] [Medline: 22468663]
39. Grittner U, Kuntsche S, Gmel G, Bloomfield K. Alcohol consumption and social inequality at the individual and country levels--results from an international study. Eur J Public Health 2013 Apr;23(2):332-339 [FREE Full text] [doi: 10.1093/eurpub/cks044] [Medline: 22562712]
40. Cutler D, Lleras-Muney A. Understanding differences in health behaviors by education. J Health Econ 2010 Jan;29(1):1-28 [FREE Full text] [doi: 10.1016/j.jhealeco.2009.10.003] [Medline: 19963292]

41. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006;8(3):e20; author reply e21 [FREE Full text] [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
42. Strecher V, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
43. Geraghty AW, Torres L, Leykin Y, Pérez-Stable EJ, Muñoz RF. Understanding attrition from international Internet health interventions: a step towards global eHealth. Health Promot Int 2013 Sep;28(3):442-452 [FREE Full text] [doi: 10.1093/heapro/das029] [Medline: 22786673]
44. Rogers E. Diffusion of Innovations. New York: Free Press; 2003.
45. Price M, Gros D, McCauley J, Gros K, Ruggiero K. Nonuse and dropout attrition for a web-based mental health intervention delivered in a post-disaster context. Psychiatry 2012;75(3):267-284 [FREE Full text] [doi: 10.1521/psyc.2012.75.3.267] [Medline: 22913502]
46. Murray E, White I, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
47. Peels D, Bolman C, Golsteijn RH, de Vries H, Mudde A, van Stralen MM, et al. Differences in reach and attrition between Web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res 2012;14(6):e179 [FREE Full text] [doi: 10.2196/jmir.2229] [Medline: 23246790]
48. Ezendam Nicole P M, Brug J, Burg J, Borsboom G, van EP, Oenema A. Differential effects of the computer-tailored FATaintPHAT programme on dietary behaviours according to sociodemographic, cognitive and home environmental factors. Public Health Nutr 2014 Feb;17(2):431-439. [doi: 10.1017/S1368980012005344] [Medline: 23257166]
49. Habibović M, Cuijpers P, Alings M, van der Voort P, Theuns D, Bouwels L, et al. Attrition and adherence in a WEB-Based Distress Management Program for Implantable Cardioverter defibrillator Patients (WEBCARE): randomized controlled trial. J Med Internet Res 2014;16(2):e52 [FREE Full text] [doi: 10.2196/jmir.2809] [Medline: 24583632]
50. Kelders S, Kok R, Ossebaard H, 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]
51. Postel G, de Haan HA, ter Huurne ED, van der Palen J, Becker S, de Jong CA. Attrition in web-based treatment for problem drinkers. J Med Internet Res 2011;13(4):e117 [FREE Full text] [doi: 10.2196/jmir.1811] [Medline: 22201703]
52. Skelton J, Irby M, Geiger A. A systematic review of satisfaction and pediatric obesity treatment: new avenues for addressing attrition. J Healthc Qual 2014;36(4):5-22 [FREE Full text] [doi: 10.1111/jhq.12003] [Medline: 23414547]
53. van der Wulp NY, Hoving C, de Vries H. A qualitative investigation of alcohol use advice during pregnancy: experiences of Dutch midwives, pregnant women and their partners. Midwifery 2013 Nov;29(11):e89-e98. [doi: 10.1016/j.midw.2012.11.014] [Medline: 23434309]
54. 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]
55. Schulz D, Kremers S, van Osch LA, Schneider F, van Adrichem MJ, de Vries H. Testing a Dutch web-based tailored lifestyle programme among adults: a study protocol. BMC Public Health 2011;11:108 [FREE Full text] [doi: 10.1186/1471-2458-11-108] [Medline: 21324181]
56. Voncken-Brewster V, Tange H, de Vries H, Nagykaldi Z, Winkens B, van der Weijden T. A randomised controlled trial testing a web-based, computer-tailored self-management intervention for people with or at risk for chronic obstructive pulmonary disease: a study protocol. BMC Public Health 2013;13:557 [FREE Full text] [doi: 10.1186/1471-2458-13-557] [Medline: 23742208]
57. Voncken-Brewster V, Tange H, de Vries H, Nagykaldi Z, Winkens B, van der Weijden T. A randomized controlled trial evaluating the effectiveness of a web-based, computer-tailored self-management intervention for people with or at risk for COPD. Int J Chron Obstruct Pulmon Dis 2015;10:1061-1073 [FREE Full text] [doi: 10.2147/COPD.S81295] [Medline: 26089656]
58. Smit E, de Vries H, Hoving C. The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses. Contemp Clin Trials 2010 May;31(3):251-258. [doi: 10.1016/j.cct.2010.03.001] [Medline: 20215047]
59. Smit E, de Vries H, Hoving C. The PAS study: a randomized controlled trial evaluating the effectiveness of a web-based multiple tailored smoking cessation programme and tailored counselling by practice nurses. Contemp Clin Trials 2010 May;31(3):251-258. [doi: 10.1016/j.cct.2010.03.001] [Medline: 20215047]
60. Stanczyk N, Bolman C, Michel JL Walthouwer, de Vries H. Study protocol of a Dutch smoking cessation e-health program. BMC Public Health 2011;11:847 [FREE Full text] [doi: 10.1186/1471-2458-11-847] [Medline: 22059446]
61. Stanczyk N, Smit E, Schulz DN, de Vries H, Bolman C, Muris JW, et al. An economic evaluation of a video- and text-based computer-tailored intervention for smoking cessation: a cost-effectiveness and cost-utility analysis of a randomized controlled trial. PLoS One 2014;9(10):e110117 [FREE Full text] [doi: 10.1371/journal.pone.0110117] [Medline: 25310007]

62. Stanczyk N, Bolman C, van Adrichem M, Candel M, Muris J, de Vries H. 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]
63. Elfeddali I, Bolman C, de Vries H. SQ4U - a computer tailored smoking relapse prevention program incorporating planning strategy assignments and multiple feedback time points after the quit-attempt: development and design protocol. Contemp Clin Trials 2012 Jan;33(1):151-158. [doi: 10.1016/j.cct.2011.09.008] [Medline: 22015574]
64. Walthouwer MJ, Oenema A, Soetens K, Lechner L, de Vries H. Systematic development of a text-driven and a video-driven web-based computer-tailored obesity prevention intervention. BMC Public Health 2013;13:978 [FREE Full text] [doi: 10.1186/1471-2458-13-978] [Medline: 24138937]
65. Walthouwer MJ, Oenema A, Soetens K, Lechner L, de Vries H. Are clusters of dietary patterns and cluster membership stable over time? Results of a longitudinal cluster analysis study. Appetite 2014 Nov;82:154-159. [doi: 10.1016/j.appet.2014.07.025] [Medline: 25064671]
66. Luijkx R, de Heus M. The educational system of the Netherlands. In: The International Standard Classification of Education. Mannheim: Mannheimer Zentrum fu¨r Europa¨ische Sozialforschung (MZES); 2008:47-75.
67. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;57:289-300. [doi: 10.2307/2346101]
68. Yekutieli D, Benjamini Y. The dontrol of the false discovery rate in multiple testing under dependency. The Annals of Statistics 2001 Aug;29(4):1165-1188. [doi: 10.1214/aos/1013699998]
69. Weinkauf M. marum. Software and programs URL: https://marum.de/Software_and_Programs.html [accessed 2015-10-03] [WebCite Cache ID 6bzOy8yh1]
70. Cumming G. Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis (Multivariate Applications Series). New York: Routledge Academic; 2012.
71. Borenstein M, Hedges LV, Rothstein H. Introduction to Meta-Analysis. Oxford: Wiley; 2009.
72. Statistisches Bundesamt. 2013. Bildungsstand URL: https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Bildungsstand/Tabellen/Bildungsabschluss.html;jsessionid=7E781B5507CB5855D3AF5C8D883FB9B3.cae3 [accessed 2015-10-03] [WebCite Cache ID 6bzMTfwTL]
73. Bierings H. Centraal Bureau voor de Statistiek. 2013 Oct 03. Onderwijsniveau bevolking gestegen URL: http://www.cbs.nl/nl-NL/menu/themas/onderwijs/publicaties/artikelen/archief/2013/2013-3905-wm.htm [accessed 2015-10-03] [WebCite Cache ID 6bzMiKF0P]
74. Huedo-Medina T, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods 2006 Jun;11(2):193-206. [doi: 10.1037/1082-989X.11.2.193] [Medline: 16784338]
75. Higgins JP, Thompson S. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
76. Cutrona SL, Mazor KM, Vieux SN, Luger TM, Volkman JE, Finney Rutten LJ. Health information-seeking on behalf of others: characteristics of "surrogate seekers". J Cancer Educ 2015 Mar;30(1):12-19. [doi: 10.1007/s13187-014-0701-3] [Medline: 24989816]
77. Jensen J, King A, Davis L, Guntzviller L. Utilization of internet technology by low-income adults: the role of health literacy, health numeracy, and computer assistance. J Aging Health 2010 Sep;22(6):804-826. [doi: 10.1177/0898264310366161] [Medline: 20495159]
78. Wangberg S, Andreassen H, Prokosch H, Santana SM, Sørensen T, Chronaki C. Relations between Internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77 [FREE Full text] [doi: 10.1093/heapro/dam039] [Medline: 18083686]
79. Wilson K, Senay I, Durantini M, Sánchez F, Hennessy M, Spring B, et al. When it comes to lifestyle recommendations, more is sometimes less: a meta-analysis of theoretical assumptions underlying the effectiveness of interventions promoting multiple behavior domain change. Psychol Bull 2015 Mar;141(2):474-509. [doi: 10.1037/a0038295] [Medline: 25528345]
80. Crutzen R, Ruiter R. Interest in behavior change interventions: a conceptual model. Eur Health Psychol 2015;17:6-11 [FREE Full text]
81. Postel M, de Haan HA, ter Huurne ED, Becker E, de Jong CA. Effectiveness of a web-based intervention for problem drinkers and reasons for dropout: randomized controlled trial. J Med Internet Res 2010;12(4):e68 [FREE Full text] [doi: 10.2196/jmir.1642] [Medline: 21163776]
82. Christensen H, Griffiths K, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
83. Kraft P, Yardley L. Current issues and new directions in Psychology and Health: What is the future of digital interventions for health behaviour change? Psychol Health 2009 Jul;24(6):615-618. [doi: 10.1080/08870440903068581] [Medline: 20205015]
84. 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]
85. McKay H, Danaher B, Seeley J, Lichtenstein E, Gau J. 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]

86. Geraghty AW, Wood A, Hyland M. Attrition from self-directed interventions: investigating the relationship between psychological predictors, intervention content and dropout from a body dissatisfaction intervention. Soc Sci Med 2010 Jul;71(1):30-37. [doi: 10.1016/j.socscimed.2010.03.007] [Medline: 20400220]
87. Kickbusch I. Health literacy: addressing the health and education divide. Health Promot Int 2001 Sep;16(3):289-297 [FREE Full text] [Medline: 11509466]
88. Paasche-Orlow M, Parker R, Gazmararian J, Nielsen-Bohlman L, Rudd R. The prevalence of limited health literacy. J Gen Intern Med 2005 Feb;20(2):175-184 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.40245.x] [Medline: 15836552]
89. Baumeister R, Heatherton T, Tice D. Losing Control: How and Why People Fail at Self-Regulation. San Diego: Academic Press; 1994.
90. Baumeister R, Muraven M, Tice D. Ego eepletion: a resource model of volition, self-regulation, and controlled processing. Social Cognition 2000 Jun;18(2):130-150. [doi: 10.1521/soco.2000.18.2.130]
91. Murray E, White I, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]
92. De Leon E, Fuentes L, Cohen J. Characterizing periodic messaging interventions across health behaviors and media: systematic review. J Med Internet Res 2014;16(3):e93 [FREE Full text] [doi: 10.2196/jmir.2837] [Medline: 24667840]
93. Fry J, Neff R. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
94. Bell A. "I think about Oprah": social class differences in sources of health information. Qual Health Res 2014 Apr;24(4):506-516. [doi: 10.1177/1049732314524637] [Medline: 24623661]
95. 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]
96. Kalichman S, Cherry C, Cain D, Weinhardt L, Benotsch E, Pope H, et al. Health information on the Internet and people living with HIV/AIDS: information evaluation and coping styles. Health Psychol 2006 Mar;25(2):205-210. [doi: 10.1037/0278-6133.25.2.205] [Medline: 16569112]
97. Elfeddali I, Bolman C, Candel Math JJ, Wiers W, de Vries H. Preventing smoking relapse via Web-based computer-tailored feedback: a randomized controlled trial. J Med Internet Res 2012;14(4):e109 [FREE Full text] [doi: 10.2196/jmir.2057] [Medline: 22903145]
98. Wienert J, Kuhlmann T. A stitch in time saves nine: things to consider when tailoring your online intervention. The Europeam Health Psychologist 2015;17:12-17 [FREE Full text]
99. Noguchi K, Albarracín D, Durantini M, Glasman L. Who participates in which health promotion programs? A meta-analysis of motivations underlying enrollment and retention in HIV-prevention interventions. Psychol Bull 2007 Nov;133(6):955-975. [doi: 10.1037/0033-2909.133.6.955] [Medline: 17967090]

Abbreviations

AEL: Alcohol-Everything within the Limit
COPD: chronic obstructive pulmonary disease
HSD: honestly significant difference
MHB: My Healthy Behavior
MYB: Master Your Breath
PAS: Personal Advice in Stopping smoking
RCT: randomized controlled trial
SQ4U: Stay Quit for You
STQ: Support to Quit
WIB: Weight in Balance

Edited by P Bamidis; submitted 17.07.15; peer-reviewed by P Krebs, P Antoniou, E Paraskevopoulos; comments to author 12.08.15; revised version received 08.09.15; accepted 22.09.15; published 07.10.15.

Please cite as:
Reinwand DA, Crutzen R, Elfeddali I, Schneider F, Schulz DN, Smit ES, Stanczyk NE, Tange H, Voncken-Brewster V, Walthouwer MJL, Hoving C, de Vries H
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
URL: http://www.jmir.org/2015/10/e228/
doi:10.2196/jmir.4941
PMID:26446779

Original Paper

Views of People With High and Low Levels of Health Literacy About a Digital Intervention to Promote Physical Activity for Diabetes: A Qualitative Study in Five Countries

Alison Rowsell[1,2*], MPhil; Ingrid Muller[1,3*], PhD; Elizabeth Murray[4], PhD, FRCGP, FRCP (Edin); Paul Little[3], FMedSci, FRCGP; Christopher D Byrne[5,6], PhD, FRCPath, FRCP; Kristin Ganahl[7], MA; Gabriele Müller[8], Dipl Ing; Sarah Gibney[9], PhD; Courtney R Lyles[10], PhD; Antonia Lucas[1], MSc; Don Nutbeam[11], PhD; Lucy Yardley[1], PhD

[1]Department of Psychology, Faculty of Social and Human Sciences, University of Southampton, Southampton, United Kingdom

[2]Faculty of Health Sciences, University of Southampton, Southampton, United Kingdom

[3]Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[4]Research Department of Primary Care and Population Health, University College London, London, United Kingdom

[5]Nutrition and Metabolism, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[6]Southampton National Institute for Health Research, Biomedical Research Centre, University Hospital Southampton, Southampton, United Kingdom

[7]Ludwig Boltzmann Institute of Health Promotion Research, Vienna, Austria

[8]Centre for Evidence-based Healthcare, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, Germany

[9]UCD School of Business, University College Dublin, Dublin, Ireland

[10]Division of General Internal Medicine at San Francisco General Hospital, University of California, San Francisco, CA, United States

[11]Office of the Vice-Chancellor, University of Southampton, Southampton, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Ingrid Muller, PhD
Department of Psychology
Faculty of Social and Human Sciences
University of Southampton
Highfield Campus
Highfield
Southampton, SO17 1BJ
United Kingdom
Phone: 44 02380 592581
Fax: 44 02380 592581
Email: I.Muller@soton.ac.uk

Abstract

Background: Low health literacy is associated with poor health-related knowledge, illness self-management, health service use, health, and survival, and thus addressing issues related to low health literacy has been highlighted as a pressing international priority.

Objective: To explore views of a digital health promotion intervention designed to be accessible to people with lower levels of health literacy, in particular examining reactions to the interactive and audiovisual elements of the intervention.

Methods: Qualitative think-aloud interviews were carried out with 65 adults with type 2 diabetes in the UK, Ireland, USA, Germany, and Austria, with purposive sampling to ensure representation of people with lower levels of health literacy. Inductive thematic analysis was used to identify common themes. We then systematically compared views in subgroups based on country, health literacy level, age, gender, and time since diagnosis.

Results: Most participants from the chosen countries expressed positive views of most elements and features of the intervention. Some interactive and audiovisual elements required modification to increase their usability and perceived credibility and relevance. There were some differences in views based on age and gender, but very few differences relating to health literacy level or time since diagnosis.

Conclusions: In general, participants found the intervention content and format accessible, appropriate, engaging, and motivating. Digital interventions can and should be designed to be accessible and engaging for people with a wide range of health literacy levels.

(J Med Internet Res 2015;17(10):e230) doi:10.2196/jmir.4999

KEYWORDS

health literacy, digital intervention, diabetes, qualitative, physical activity

Introduction

Addressing the problem of low health literacy has been highlighted as a pressing international priority, as it is associated with poor health-related knowledge, illness self-management, health service use, health, and survival [1]. The concept of health literacy has evolved to refer to "the knowledge, motivation and competences to access, understand and apply health information" [2]. Health literacy levels tend to be lower in those with less education, lower incomes, and minority ethnics groups [1]. Health literacy is a product of the interaction between the individual and his or her environment, which includes the health care resources available to him or her; designing health care materials to be accessible and easily comprehensible can reduce the literacy burden, and improve health literacy by helping people to understand and implement health-related advice [3].

Many people have difficulty accessing face-to-face diabetes self-management education, due to barriers such as work, caring responsibilities, disability, cost, and lack of transport [4], and these problems are more common among those with less education [5]. The rapid growth in delivery of health promotion and health care by means of digital interventions offers one possible solution to this challenge; digital interventions can be accessed conveniently at home and have the potential for wide reach at low cost, and so could reduce health disparities [6]. However, there is a risk that digital interventions could increase health inequalities due to a "digital divide" in both access to the Internet and confidence and skills to use the Internet for self-management of health [7]. Internet access is now increasing rapidly among all social groups, but low literacy levels may continue to pose barriers to understanding and applying online information. For example, one study of diabetes websites found that 86.9% of materials would be too difficult to read for an average adult [8], and from early studies of Internet-delivered support for people with diabetes there is evidence of lower usage by those with lower income and education [9,10].

Interventions to reduce the literacy burden and improve health literacy have included using simple language; information presented in audio, audiovisual, or pictorial formats; (when Internet delivered) employing tailoring of content to individual's needs; and other forms of interactivity. Reviews of the effectiveness of such interventions for general public and mixed patient populations [11-15] and for diabetes [16-18] suggest that these approaches show promise for some outcomes, but that overall the evidence is weak and inconclusive. Few studies have been theory based, and it remains unclear exactly which elements of such interventions improve which outcomes. The authors of these studies call for more research to identify the effective components and mechanisms of interventions to improve health literacy and to examine whether these differ for people with different levels of health literacy [11,14,15]. Empirical evidence that addresses the latter question is crucial to inform those delivering health care about whether it is necessary to create different versions of online interventions to engage different sectors of the population; such evidence will consequently have wide-ranging practical and resource implications.

As part of the Diabetes Literacy project [18], our aim was to address this evidence gap by examining how people with varying levels of health literacy viewed features of a digital intervention we developed to explore how best to increase accessibility and acceptability of health promotion resources for people with lower levels of health literacy. Our primary aim was therefore to examine whether our website design was acceptable and engaging for people with differing levels of health literacy. Because digital delivery permits interactive and audiovisual presentation of advice, that could potentially help to overcome difficulties due to low literacy, we were also particularly interested in exploring participants' views of these features of the intervention.

Methods

Intervention

The intervention was a website designed to motivate people with diabetes to increase their levels of physical activity. The interactive features of the website comprised tailoring of images and advice based on user responses to questions (eg, about the user's age, concerns about physical activity, current activity levels), a quiz, and a physical activity planner. The audiovisual features included positive images and audiovisual sequences illustrating lifestyle physical activities. We followed established good practice for designing written medical information and making it accessible to people of all literacy levels [11,19], and used design principles that have been shown to increase accessibility of websites for people with cognitive impairment and limited computer literacy [20]. See Textbox 1 for further details of the elements of the intervention, and Figure 1 for example screenshots.

The intervention was developed in consultation with an expert panel of patient representatives, clinicians, and behavioral scientists. We employed the person-based approach to intervention development [21], which grounds intervention design in a rigorous, in-depth understanding of the psychosocial context of the target user population. A key element of the person-based approach to intervention development is to use iterative, inductive qualitative research to explore users' views of the intervention and then modify the intervention to optimize

acceptability and engagement. Consequently, small changes to the website were made throughout data collection. These included, for example, changing the format of the physical activity planner to improve usability, changing wording to improve comprehension, and substituting images that were disliked. The intervention was developed using *LifeGuide* software, a platform for developing online behavior change interventions that allows researchers to easily translate and modify the intervention [22]. The website was initially developed in English for evaluation in the UK, Ireland, and USA, and then translated for evaluation in Germany and Austria (all text of the translated versions was checked for accuracy by the researchers in these countries).

Textbox 1. Elements and design features of the Healthy Living with Diabetes intervention.

Interactive elements of the intervention

- Delivery of information about the health benefits of physical activity for people with diabetes (and health risks of inactivity) in the form of an interactive "fun" quiz.
- Tailored advice in response to questions about current physical activity level and concerns about physical activity (eg, barriers such as cost, health problems).
- Tailoring of images based on users' reported age.
- A physical activity planner that enabled users to create personal plans for increasing physical activity, building on their current activities.

Audiovisual elements of the intervention

- Positive visual images throughout the intervention.
- Audiovisual sequences modeling people undertaking a range of physical activities: these comprised a narrative illustrated by a sequence of photographs with a voice-over, and had an informal tone intended to suggest real-life scenarios.
- Option to access details of study aims and procedures in audio format.

Good practice design features of the intervention

- To maximize accessibility to those with low levels of literacy, text as short as possible, suitable for reading age of 12.
- To maximize accessibility to those with low levels of computer literacy, simple page layout and navigation.
- Advice appropriate for those with low current levels of physical activity, limited time, motivation and resources, and health problems (ie, promoted gradual increase in preferred lifestyle-compatible activities).

Figure 1. Screenshots illustrating elements of the Healthy Living with Diabetes website. (A) Example of positive visual image on welcome page. (B) Example of quiz feedback with positive visual image. (C) Example of audiovisual sequence. (D) Example of interactive physical activity planner.

Design, Participants, and Procedure

Qualitative interviews were carried out with adults with type 2 diabetes in the UK, Ireland, USA, Germany, and Austria. Full details of all study procedures, participants, and analyses are published online [23]. The UK sample was used as the core group for intervention development and evaluation, and data collection in the other countries was used to explore whether views differed in different countries and settings. Participants were recruited from primary care, community settings, and diabetes support groups, using advertisements, letters, and personal invitations; for full details see [23]. We purposively sampled to include people with lower levels of health literacy by recruiting from areas and clinics with high levels of socially deprived patients. We also used diversity sampling to ensure that we had a broad balance of both genders, and different ages and time since diagnosis.

Qualitative think-aloud methods [24] were used to gain a thorough, in-depth understanding of user experiences, perceptions, and views of the intervention, followed by a semistructured interview, allowing specific domains of user experience to be explored. Experienced researchers who received standardized training on conducting think-aloud interviews interviewed participants in their own language in each country, in a variety of settings in clinics and the community. Participants also completed a brief questionnaire to measure age, gender, time since diagnosis, and health literacy. Health literacy was measured using the validated single item "how often do you have problems learning about your condition because of difficulty understanding written information?"[25], allowing participants to be categorized as having high, intermediate, or low levels of health literacy. To enable us to validate this single-item, categorical measure of health literacy, we examined its correlation with a more comprehensive measure of all dimensions of health literacy, the 16-item version of the European Health Literacy Survey Questionnaire (HLS-EU-16) [26]; this was completed by all participants after the interviews.

Analysis

Audio recordings of interviews were transcribed verbatim, translated into English where necessary, and checked for accuracy by the researcher who carried out the interview. Few substantial differences in views between countries were identified (see [23] for details), allowing the data to be pooled to give a larger sample for the remaining comparisons. Inductive thematic analysis was used to identify recurring themes through close examination of the data [27,28]. We also used analytic techniques from grounded theory to increase the rigor of our analysis, including line-by-line coding and constant comparison [29]. First, interview transcripts were read and reread to ensure a high level of familiarity with the data before line-by-line coding of the initial 3 interviews. A provisional coding manual (see [23] for details) was then created to define emerging codes and themes before these codes were applied to the remaining

transcripts. The coding manual was developed iteratively and revised throughout the coding process to ensure the codes adequately reflected the data. The coding manual was discussed and agreed upon by core members of the research team (AR, IM, and LY) at various stages of the coding process, and inter-rater agreement between AR and IM was obtained for all the final coded data.

The relatively large sample size (for qualitative research) allowed us to carry out a second stage of analysis to explore whether any differences could be identified between specific subgroups of participants. We did this by creating tables to systematically compare the occurrence and content of themes across subgroups (see). First, country comparisons were carried out to examine whether there were differences in the data between countries that might preclude pooling the data across countries for analysis. Only the data from the German sample appeared systematically different, largely in terms of less views being expressed on most topics. This finding may have been related to deviations from the interview protocol evident in the transcripts (eg, a less open-ended interviewing style than in other countries, omission of substantial parts of the interview schedule). In view of this difference, the 6 German interviews were excluded from the subsequent subgroup analyses. Further subgroup analyses on the remaining pooled data comprised comparisons based on level of health literacy (low, intermediate, or high), age (under or over 60 years), gender, and time since diagnosis (less or more than 5 years).The cutoff points were chosen so as to create roughly equal groups (to maximize the sample size in each) that were likely to differ in terms of experiences of diabetes (based on the research team's clinical experience of diabetes patients). Because of the small size of the subgroups and large number of comparisons, to avoid overinterpretation of minor differences we adopted a criterion of only reporting comparisons where the differences observed between the groups compared were based on at least five people (between country comparisons) or 30% (n=19) of participants (pooled data), or suggested a consistent trend from high to low health literacy (even if not quite reaching our criterion within the specific subgroups).

Results

Participant Characteristics

In total, 65 participants were interviewed for this study, comprising 35 from the UK, 8 from USA, Ireland and Austria, and 6 from Germany (Table 1). There were 37 (57%) men and 28 (43%) women, with a mean age of 62 years (range 37–79), and 40 (62%) participants had been diagnosed with type 2 diabetes for 5 years or more. As classified by our single-item measure, 38 (58%) participants had high health literacy, 18 (28%) had intermediate health literacy, and 8 (12%) had low health literacy (1 was "unknown"). The single-item measure was highly correlated with the HLS-EU-16 in a subsample sample (r=.65, n=35, P<.001), providing reassurance that our single-item categorization assessing literacy problems in a health context was associated with health literacy problems in a range of other domains. Participant characteristics were similar across all countries, with the exception of Germany (all German participants had high levels of health literacy and all but 1 person had diabetes for over 5 years).

Table 1. Participant characteristics.

Participant characteristics		Country					
		UK	USA	Ireland	Austria	Germany	Overall
Age (in years)							
	Mean	58	57	66	60	67	62
	Range	44-75	49-64	37-77	42-79	48-77	37-79
	≥60 years	18	3	6	3	4	34
	<60 years	17	5	2	5	2	31
Time since diagnosis							
	<5 years	14	4	4	2	1	25
	≥5 years	21	4	4	6	5	40
Gender							
	Male	18	3	6	6	4	37
	Female	17	5	2	2	2	28
Health literacy level							
	High	16	4	6	7	5	38
	Intermediate	13	2	2	0	1	18
	Low	6	1	0	1	0	8
	Not known	0	1	0	0	0	1
Ethnicity							
	White/Caucasian	32	2	7	8	6	55
	Black/African/Caribbean	2	2	0	0	0	4
	Asian	1	0	1	0	0	2
	Other (mixed)	0	4	0	0	0	4
Total sample		35	8	8	8	6	65

Themes Emerging From Pooled Analysis

Thematic analysis of the data pooled across countries generated 40 codes, which we organized into 3 main themes, namely, (1) general reactions to website content and format; (2) reactions to interactive features; and (3) reactions to audiovisual features.

General Reactions to Website Content and Format

The majority of users described the level of information in the intervention as appropriate and easy to understand, without being patronizing (Table 2). The website was also described as accessible and as more user friendly than other resources they had encountered.

> *I like the whole thing. It didn't blind you with science; it didn't treat you as a total idiot.* [UK, male, under 60 years old, under 5 years' diagnosis, high HL]
>
> *It's written in a way that's not—because I mean I'm not good at reading and stuff like that—but it's written in a way that I can understand. You know, sometimes you look at, whether it's books or websites or whatever, and sometimes you're reading and you think "What are they talking about?"* [UK, female, under 60 years old, over 5 years' diagnosis, low HL]
>
> *The thing is that all the knowledge I received during hours of lectures [at "wellness clinics"] is briefly put together here, so really, really well done for this!* [Austria, male, under 60 years old, under 5 years' diagnosis, low HL]
>
> *No, it's very straight forward and compact. What you need to know is there.* [Austria, male, under 60 years old, over 5 years' diagnosis, high HL]

Participants frequently spoke appreciatively of acquiring new information through the website and the positive framing of health information was described as encouraging.

> *Yeah, I know they say physical, mental activity can, can help you stay healthy [reads to self], "twice," woah, "twice as likely!" Get out, really? I don't remember ever hearing that one.* [USA, female, over 60 years old, over 5 years' diagnosis, high HL]
>
> *To find out that Alzheimer's thing, yes that was quite shocking, a sit up and take notice moment.* [UK, male, under 60 years old, over 5 years' diagnosis, intermediate HL]

A number of people discussed their intention to increase their physical activity as a direct result of what they had learned or seen in the website.

Interesting, I did not know that—so it's almost like, if I learn nothing else from this survey, I need to start my physical activity regime. [USA, male, over 60 years old, under 5 years' diagnosis, intermediate HL]

Specific intervention features described as motivating included the planner and audiovisual sequences as well as learning new health information through the quiz. Details of how participants viewed these features are given below.

Table 2. Reactions to website content and format.

Subtheme	Content
Information novelty	Almost all participants mentioned learning new, often surprising information (particularly about the benefits of physical activity for preventing Alzheimer's disease and for healthy liver function)
Level of website advice	Most participants felt that the advice was delivered at the right level—easy to understand but not patronizing
Views of website advice	The vast majority of participants endorsed the advice given by the website.
	Many were surprised and some skeptical about the information that physical activity is more important than controlling blood sugar levels for preventing complications from diabetes.
	The humor in the website was mainly appreciated by the minority who commented on it.
Views on website appearance	Most participants commented that they found the website clear and easy to navigate.
	Some expressed a desire for greater simplicity in presentation and less text.
Effects of website on motivation	Most participants (though not all) found the website generally motivating.
	All elements of website content were described by some participants as motivating them to engage in greater physical activity.

Reactions to the Interactive Features

Many participants commented that the interactive features of the website were engaging and motivating (Table 3), although some chose to skip sections that did not appeal to them. Most were positive about the quiz section, describing it as fun, relevant, helpful, interesting, and a preferable way of learning new information.

I'm getting really curious now to see the answers [both laugh], it's lovely, it's not boring, the whole Web thing is very good. It's very interesting. [Ireland, female, over 60 years old, over 5 years' diagnosis, high HL]

I liked the quizzes, I liked that—you know—it's nice to have something you can use, interact with and join in with. [UK, female, under 60 years old, under 5 years' diagnosis, low HL]

If I was doing this on my own I would skip this bit [quiz section], I'm bored with that now, because it's treating me like a child. I want the information but I don't like the way it's given to you. [UK, male, under 60 years old, under 5 years' diagnosis, low HL]

Participants described valuing the instant tailored feedback, and often commented on encouragement provided by the positive framing of this advice.

Woo hoo—"doing enough physical activity to keep healthy"—now says star pupil actually, that's a bonus. [UK, female, over 60 years old, over 5 years' diagnosis, low HL]

Well done, you got it right, you got something right. See, they love me. Finally! [laughs] They love me. [USA, female, over 60 years old, under 5 years' diagnosis, unknown HL]

However, participants had more mixed responses to the activity planner, some finding it cumbersome and difficult to complete, particularly (but not exclusively) before modifications were made to simplify it. For example, the original planner required participants to specify the activity type and the amount of time spent on the activity on the same page. Many participants found this difficult to navigate, and the planner was subsequently separated onto 2 pages.

Table 3. Reactions to the interactive features.

Subtheme	Content
Views of the interactive quiz	Most comments on the quiz were that it was enjoyable and informative.
	Some found it irritating, and disliked the humorous "trick" questions (ie, that physical activity would not improve hearing or alcohol consumption).
Views of tailored feedback	Most comments on the feedback were that participants appreciated getting immediate, positive feedback.
	Participants noted that actively engaging with the website kept their attention.
Views on the interactive activity planner	Some participants found using the planner difficult, due mainly to uncertainty about (1) how to estimate their activity level and (2) how to enter data into the planner (before simplification of the planner)

Reactions to the Audiovisual Features

The vast majority of participants found the images acceptable or liked them overall (Table 4).

> *These pictures are quite nice as well, you know, they are happy pictures—photographs of elderly people, probably my age, dancing with umbrellas.* [Ireland, male, over 60 years old, over 5 years' diagnosis, high HL]

A few participants mentioned disliking certain images, although the images that people disliked varied. For example, some younger participants could not relate to images of older people; consequently, the intervention was modified to tailor all images by age.

Most participants also enjoyed the audiovisual sequences, appreciating the informal style and relating positively to the stories and example activities these sequences narrated.

> *The videos were really, really good...informative.* [USA, male, over 60 years old, under 5 years' diagnosis, low HL]

> *It was good, straight to the point, because it's like, not everybody likes going to the gym and that. It's too expensive, too hot to try to stay and that, and the person that was speaking was really clear.* [UK, male, under 60 years old, over 5 years' diagnosis, low HL]

Some participants commented positively on the use of humor in the main audiovisual sequence (which showed pictures of snowboarding and deep sea diving as examples of impractical forms of exercise before suggesting walking as more feasible) but some found it confusing or inappropriate. In addition, some people disliked particular scenarios that they found unconvincing or irrelevant:

> *I don't like it at all. It doesn't fit the pictures, there's an old man who says he's picking up his children. He says his day is hectic yet he can go shopping before work and during lunch. It doesn't fit.* [Germany, male, under 60 years old, over 5 years' diagnosis, high HL]

Some of the scenarios therefore required modification for different cultures or age groups; for example, a scenario that suggested playing football was changed to baseball for US participants.

Table 4. Reactions to the audiovisual features.

Subtheme	Content
Views of visual images in the website	Overall, use of images in website appreciated as positive, attractive.
	Images of walking and family activities generally appreciated, medical illustrations mainly well received.
	Reactions to some images mixed (prior to modification), for example, images of older people, wheelchair activity, alcohol.
Views of audiovisual sequences in the website	Most participants were very positive about the audiovisual sequences, most (though not all) liking the informal style, and finding them engaging, funny.
	Most people enjoyed relating to stories that they saw as realistic, helpful examples of lifestyle activity.
	Negative comments were often based on seeing specific content (prior to modification) as unrealistic or irrelevant to the participant's own situation (eg, due to cultural or lifestyle differences or activity preferences).

Subgroup Comparisons

We found few differences in how people with high, intermediate, or low levels of health literacy viewed the website. The only differences that emerged were that people with higher health literacy were most likely to comment that they found the website content easy to understand, mention features of the website they found motivating, discuss how the interactivity maintained their attention, and express appreciation of the visual images. There were also few systematic differences in views linked to country, although there were some culture-specific preferences (as noted earlier); for example, the Austrian participants tended to dislike the audiovisual sequences, commenting negatively on the use of speakers with German accents.

With regard to other subgroup comparisons, women were more likely than men to voice their intention to be more physically active as a result of viewing the website. Women tended to express much more positive views of the quiz than men did; some men found it tedious or irrelevant. Women were also more positive about the audiovisual sequences, describing them as engaging and discussing relating to the characters and stories. Women, however, were less technically confident in completing the interactive planner. There were age differences in the activities preferred; those aged over 60 mainly intended to do more walking, whereas younger participants also spoke about cycling and swimming. Participants over 60 years of age and those who had been diagnosed for longer were more positive about the audiovisual sequences (particularly the walking stories) and the idea of exercising at home. This age group was also more likely to mention that they found the level of information in the website straightforward, helpful, or pitched at the correct level.

Discussion

Preliminary Findings

The main aim of this study was to investigate whether it is possible to design a digital intervention that is acceptable and engaging for people with varying levels of health literacy. Our findings are encouraging; most participants from most countries expressed positive views about most elements of our digital intervention. Very few people found the accessible format

patronizing or the information provided inadequate or inappropriate. Despite concerns frequently expressed about whether it is possible to appeal to a wide and diverse target population, it is not unusual to find that materials designed to be accessible to those with lower levels of literacy and health literacy are also liked by those with higher literacy levels [30-32].

Although most participants had mainly positive reactions to the interactive and audiovisual presentation of advice, it was necessary to iteratively modify these based on participant feedback to optimize acceptability and feasibility. There were also differences in reactions to these elements of the intervention relating to age, gender, and cultural context. We were only able to evaluate views of a very limited set of digital materials, and it is quite possible that the views expressed were specific to these particular resources. For example, reactions to professionally produced videos of patients or actors might be entirely different from reactions to our amateur-style audiovisual sequences. Nevertheless, some previous research has also noted that it can be difficult to produce audiovisual narratives that precisely match the psychosocial context of all users [33]. It appears that audiovisual materials may require particularly careful development, with attention to the sociocultural context of target users, to ensure that they are perceived as convincing and relevant [21,34].

Limitations

The method we used for comparing views across subgroups is unusual; it can be regarded as an extension of the constant comparison technique employed in grounded theory, but with more systematic and explicit assessment of the frequency with which views were expressed in different subgroups. Because these subgroups were too small to permit reliable quantitative evaluation, it is not possible to interpret the trends observed in our data as definitive evidence for the presence or absence of group differences. Nevertheless, the consistent absence of any variations in views clearly associated with health literacy and time since diagnosis, despite clear differences based on age and gender, suggests the latter may have had a more important influence on reactions to the intervention. However, we were unable to recruit many participants with the lowest levels of health literacy, and the sample size in each country was small. Consequently, this study was only able to identify substantial differences in views due to country or very low health literacy (ie, views that would be expressed by most people in these subgroups). Most participants were also white, and these findings may therefore not be generalizable to people from other ethnic backgrounds. A further limitation in our ability to fully investigate the perspective of users with lower levels of health literacy is that these participants provided fewer comments about the website and our analysis of the articulated views of users was unable to capture nonverbal indications of accessibility or engagement barriers, such as pauses and silences. Further research into the views of people with very low health literacy is required.

It is important to be aware that, despite our best efforts to encourage participants to freely express negative views, some participants may have been reluctant to do so. For example, it is possible that women did not find the intervention more engaging than men, but were less willing to express negative views of it. It is also important to remember that for the purpose of health promotion it is necessary but not sufficient for an intervention to be acceptable and engaging—it must also be effective in achieving the intended behavioral outcomes. For this reason, we are now undertaking a large trial to test the effectiveness of our digital intervention for promoting health literacy improvement (in knowledge, understanding, and self-efficacy) and behavioral intentions with regard to increased physical activity.

Conclusion

Most participants found the intervention generally acceptable and engaging. Surprisingly, reactions were similar—and equally positive—among those with higher and lower levels of health literacy. This finding has importance for the reach and cost effectiveness of digital health care, because it suggests that it may not be necessary to develop multiple versions of interventions for people with differing levels of health literacy. However, our sample did not include people with the very lowest levels of health literacy, who may well require different online or offline interventions.

Some marked variations in preferences for how advice was presented did emerge; for example, many people particularly appreciated the interactive quizzes and audiovisual sequences, whereas a minority strongly disliked them. These variations in preferences were partly linked to age, gender, and culture but were not closely mapped onto demographic and psychosocial characteristics, making it difficult to prescribe exactly what format should be used for which population subgroup. This suggests that perhaps a major benefit of Internet delivery of health promotion is that it is possible to offer recipients a choice of formats, allowing them to self-select those that they find most accessible, attractive, and useful. This approach permits users to engage in what has been termed “self-tailoring” [35], in contrast to the pre-emptive tailoring to major preferences in target groups that intervention developers must use when constrained by the page limits of printed materials. Offering users a choice of how they engage with interventions has the potential to promote autonomous motivation [21] and is consistent with the way in which people are accustomed to using the Internet. However, conventional tailoring may still be required to ensure that users are not presented with material or elements that they find so alienating or demotivating that they simply cease using the intervention—for example, the images of older people that younger participants could not relate to in our study, or the German accents that our Austrian participants found off-putting.

Practice Implication

These findings have clear implications for those who develop health-related websites and digital interventions; these can and should be designed to be accessible and engaging for people with a wide range of levels of health literacy, to help to overcome the “digital divide” and reduce health inequalities. This can be achieved by following the established good design principles we drew on and then using findings from iterative qualitative research to maximize the perceived relevance,

credibility, and feasibility of the intervention for different members of the target population. Incorporating interactive and audiovisual elements to increase interest and engagement may also be useful but requires careful development to ensure that they are appropriate for people of both genders, different ages, and from different cultures. Finally, with regard to international dissemination of digital interventions, our findings indicate that the acceptability of interventions is likely to be similar across different countries but can also be improved by making modifications on the basis of feedback from interviews to increase the perceived relevance to the specific cultural context.

Acknowledgments

This research was part of the Diabetes Literacy project supported by a grant of the European Commission (Grant No FP7-Health-2012-Innovation-1/306186). The authors would like to thank the experts and members of the Diabetes Literacy Study Group who contributed by commenting on the intervention content, design, and implementation, in particular Robyn Dixon, Paul Roderick, Deborah Morrison, Richard Holt, Wendy Lowe, Pal Kingshuk, Scott Lloyd, Brian Williams, and Kath Barnard. We would especially like to thank our two lay advisors, Andy Hamson and Wendy Heath, and all of the people who contributed their time and efforts to the audiovisual content of the website. We are also grateful to those who helped translate and transcribe the data. CDB is supported in part by the Southampton National Institute for Health Research Biomedical Research Centre.

Conflicts of Interest

None declared.

References

1. Greenhalgh T. Health literacy: Towards system level solutions. BMJ 2015 Feb 24;350:h1026. [doi: 10.1136/bmj.h1026]
2. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: A systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
3. Nutbeam D. The evolving concept of health literacy. Soc Sci Med 2008 Dec;67(12):2072-2078. [doi: 10.1016/j.socscimed.2008.09.050] [Medline: 18952344]
4. Pereira K, Phillips B, Johnson C, Vorderstrasse A. Internet delivered diabetes self-management education: A review. Diabetes Technol Ther 2015 Jan;17(1):55-63. [doi: 10.1089/dia.2014.0155] [Medline: 25238257]
5. Thoolen B, de Ridder RD, Bensing J, Gorter K, Rutten G. Who participates in diabetes self-management interventions?: Issues of recruitment and retainment. Diabetes Educ 2007;33(3):465-474. [doi: 10.1177/0145721707301491] [Medline: 17570877]
6. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res 2010;12(5):e60 [FREE Full text] [doi: 10.2196/jmir.1463] [Medline: 21169162]
7. McAuley A. Digital health interventions: Widening access or widening inequalities? Public Health 2014 Dec;128(12):1118-1120. [doi: 10.1016/j.puhe.2014.10.008] [Medline: 25458115]
8. Kusec S, Brborovic O, Schillinger D. Diabetes websites accredited by the Health On the Net Foundation Code of Conduct: Readable or not? Stud Health Technol Inform 2003;95:655-660. [Medline: 14664062]
9. Sarkar U, Karter A, Liu J, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: Health literacy and the use of an internet-based patient portal in an integrated health system—Results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2:183-196 [FREE Full text] [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]
10. Glasgow R, Strycker L, Kurz D, Faber A, Bell H, Dickman JM, et al. Recruitment for an internet-based diabetes self-management program: Scientific and ethical implications. Ann Behav Med 2010 Aug;40(1):40-48. [doi: 10.1007/s12160-010-9189-1] [Medline: 20411443]
11. Clement S, Ibrahim S, Crichton N, Wolf M, Rowlands G. Complex interventions to improve the health of people with limited literacy: A systematic review. Patient Educ Couns 2009 Jun;75(3):340-351. [doi: 10.1016/j.pec.2009.01.008] [Medline: 19261426]
12. Ryan R, Prictor M, McLaughlin K, Hill S. Audio-visual presentation of information for informed consent for participation in clinical trials. Cochrane Database Syst Rev 2008(1):CD003717. [doi: 10.1002/14651858.CD003717.pub2] [Medline: 18254029]
13. Mackert M, Champlin S, Holton A, Muñoz I, Damásio M. eHealth and health literacy: A research methodology review. J Comput-Mediat Comm 2014 Apr 12;19(3):516-528. [doi: 10.1111/jcc4.12044]
14. Jacobs R, Lou J, Ownby R, Caballero J. A systematic review of eHealth interventions to improve health literacy. Health Informatics J 2014 Jun 10:1 (forthcoming)(forthcoming). [doi: 10.1177/1460458214534092] [Medline: 24916567]
15. Barry M, D'Eath M, Sixsmith J. Interventions for improving population health literacy: Insights from a rapid review of the evidence. J Health Commun 2013;18(12):1507-1522. [doi: 10.1080/10810730.2013.840699] [Medline: 24298885]

16. Welch G, Shayne R. Interactive behavioral technologies and diabetes self-management support: Recent research findings from clinical trials. Curr Diab Rep 2006 Apr;6(2):130-156. [Medline: 16542624]
17. Glasgow R. Interactive media for diabetes self-management: Issues in maximizing public health impact. Med Decis Making 2010;30(6):745-758. [doi: 10.1177/0272989X10385845] [Medline: 21183760]
18. Fransen M, von WC, Essink-Bot M. Diabetes self-management in patients with low health literacy: Ordering findings from literature in a health literacy framework. Patient Educ Couns 2012 Jul;88(1):44-53. [doi: 10.1016/j.pec.2011.11.015] [Medline: 22196986]
19. Hirsh D, Clerehan R, Staples M, Osborne R, Buchbinder R. Patient assessment of medication information leaflets and validation of the Evaluative Linguistic Framework (ELF). Patient Educ Couns 2009 Nov;77(2):248-254. [doi: 10.1016/j.pec.2009.03.011] [Medline: 19375267]
20. Rotondi A, Eack S, Hanusa B, Spring M, Haas G. Critical design elements of e-health applications for users with severe mental illness: Singular focus, simple architecture, prominent contents, explicit navigation, and inclusive hyperlinks. Schizophr Bull 2015 Mar;41(2):440-448. [doi: 10.1093/schbul/sbt194] [Medline: 24375458]
21. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: Application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
22. Williams S, Yardley L, Wills G. A qualitative case study of LifeGuide: users' experiences of software for developing Internet-based behaviour change interventions. Health Informatics J 2013 Mar;19(1):61-75. [doi: 10.1177/1460458212458915] [Medline: 23486826]
23. Rowsell A, Muller I, Murray E, Little P, Nutbeam D, Yardley L. Patient Views and Experiences of Web-Based Support for Diabetes Self-Management Designed for People with Different Levels of Health Literacy. Southampton, UK: University of Southampton; 2015. URL: http://www.diabetesliteracy.eu/system/files/documents/DiabetesLiteracy_D8.11.pdf [accessed 2015-07-26] [WebCite Cache ID 6aJnOpFz6]
24. van den Haak MJ, de Jong M, Schellens P. Evaluation of an informational web site: Three variants of the think-aloud method compared. Tech Commun 2007;54(1):58-71.
25. Chew L, Bradley K, Boyko E. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
26. Sørensen K, Van den Broucke S, Pelikan J, Fullam J, Doyle G, Slonska Z, HLS-EU Consortium. Measuring health literacy in populations: Illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013;13:948 [FREE Full text] [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3:77-101.
28. Joffe H, Yardley L. Content and thematic analysis. In: Marks D, Yardley L, editors. Research Methods for Clinical and Health Psychology. London, UK: Sage; 2004:56-68.
29. Charmaz K, Henwood K. Grounded theory. In: The Sage Handbook of Qualitative Research in Psychology. London, UK: Sage; 2008:240-260.
30. Longo D, Schubert S, Wright B, LeMaster J, Williams C, Clore J. Health information seeking, receipt, and use in diabetes self-management. Ann Fam Med 2010;8(4):334-340 [FREE Full text] [doi: 10.1370/afm.1115] [Medline: 20644188]
31. Smith S, Trevena L, Nutbeam D, Barratt A, McCaffery K. Information needs and preferences of low and high literacy consumers for decisions about colorectal cancer screening: Utilizing a linguistic model. Health Expect 2008 Jun;11(2):123-136. [doi: 10.1111/j.1369-7625.2008.00489.x] [Medline: 18494957]
32. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. Understanding reactions to an internet-delivered health-care intervention: Accommodating user preferences for information provision. BMC Med Inform Decis Mak 2010;10:52 [FREE Full text] [doi: 10.1186/1472-6947-10-52] [Medline: 20849599]
33. Ezendam NP, Alpay LL, Rövekamp TA, Toussaint PJ. Experimenting with case-based reasoning to present educative health information on the Internet: The example of SeniorGezond. Stud Health Technol Inform 2005;116:867-872. [Medline: 16160367]
34. Mackert M, Kahlor LA, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: Addressing the obesity epidemic. Telemed J E Health 2009 Sep;15(7):672-677 [FREE Full text] [doi: 10.1089/tmj.2009.0012] [Medline: 19694596]
35. Lorig K, Holman H. Self-management education: History, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]

Edited by G Eysenbach; submitted 02.08.15; peer-reviewed by A Rotondi, D Sims; comments to author 27.08.15; revised version received 07.09.15; accepted 22.09.15; published 12.10.15.

Please cite as:
Rowsell A, Muller I, Murray E, Little P, Byrne CD, Ganahl K, Müller G, Gibney S, Lyles CR, Lucas A, Nutbeam D, Yardley L
Views of People With High and Low Levels of Health Literacy About a Digital Intervention to Promote Physical Activity for Diabetes: A Qualitative Study in Five Countries
J Med Internet Res 2015;17(10):e230
URL: http://www.jmir.org/2015/10/e230/
doi:10.2196/jmir.4999
PMID:26459743

Original Paper

A Web-Based Self-Help Intervention With and Without Chat Counseling to Reduce Cannabis Use in Problematic Cannabis Users: Three-Arm Randomized Controlled Trial

Michael P Schaub[1], PhD; Andreas Wenger[1], MSc; Oliver Berg[2], MD; Thilo Beck[2], MD; Lars Stark[2], MA; Eveline Buehler[1], MSc; Severin Haug[1], PhD

[1]Swiss Research Institute for Public Health and Addiction (ISGF), associated to the University of Zurich and World Health Organization Collaborating Center, Zurich, Switzerland

[2]Arud Center for Addiction Medicine, Zurich, Switzerland

Corresponding Author:
Michael P Schaub, PhD
Swiss Research Institute for Public Health and Addiction (ISGF)
associated to the University of Zurich and World Health Organization Collaborating Center
Konradstrasse 32
Zurich, 8031
Switzerland
Phone: 41 444481165
Fax: 41 444481170
Email: michael.schaub@isgf.uzh.ch

Abstract

Background: After alcohol and tobacco, cannabis is the most widely used psychoactive substance in many countries worldwide. Although approximately one in ten users develops serious problems of dependency, only a minority attend outpatient addiction counseling centers. A Web-based intervention could potentially reach those users who hesitate to approach such treatment centers.

Objective: To test the efficacy of a Web-based self-help intervention with and without chat counseling—Can Reduce—in reducing the cannabis use of problematic cannabis users as an alternative to outpatient treatment services.

Methods: Altogether, 436 participants were recruited by various online and offline media for the Web-based trial. A total of 308 of these were eligible for study participation and were randomly allocated in an unblinded manner to either self-help with chat (n=114), self-help without chat (n=101), or a waiting list control group (n=93). The fully automated self-help intervention consisted of eight modules designed to reduce cannabis use, and was based on the principles of motivational interviewing, self-control practices, and methods of cognitive behavioral therapy. Additional individual chat counseling sessions were based on the same therapeutic principles. The sessions were conducted by trained counselors and addressed participants' personal problems. The main outcomes were the frequency (number of days) and quantity of cannabis use (number of standardized joints) per week, as entered into the consumption diary at baseline and at the 3-month follow-up. Secondary outcomes included self-reported symptoms of cannabis use disorder, severity of cannabis dependence, risky alcohol use, and mental health symptoms. Intervention participation and retention were extracted from the user progress data and the consumption diary, respectively.

Results: Can Reduce participants were older (U=2.296, *P*=.02) and reported a greater number of cannabis use days at baseline than patients who entered outpatient treatment with cannabis as their main problem substance (data from the Swiss treatment demand monitoring statistics were used; chi-square [df 2]=4.0, *P*=.046). Participants in the self-help with chat study arm completed a mean of 3.2 modules and 27 out of 114 (23.7%) of the participants received at least one chat session. Participants in the self-help without chat study arm completed similar numbers of self-help modules. A total of 117 of 308 participants (38.0%) completed the 3-month follow-up assessment. The change in the mean number of cannabis use days per week at 3 months differed between self-help without chat (mean change 0.7, SD -0.2) and self-help with chat (mean change 1.4, SD -0.5; beta=-0.75, SE=0.32, t=-2.39, *P*=.02, d=0.34, 95% CI 0.07-0.61), as well as between self-help with chat and waiting list (mean change 1.0, SD -0.8; beta=0.70, SE=0.32, t=2.16, *P*=.03, d=0.20, 95% CI -0.07 to 0.47). However, there were no differences between self-help without chat and waiting list (beta=-0.05, SE=0.33, t=-0.16, *P*=.87, d=-0.14, 95% CI -0.43 to 0.14). Self-reported abstinence was significantly different in the self-help without chat study arm (2.0%) than in the self-help with chat study arm (8.8%; beta=-1.56,

SE=0.79, *P*=.05, odds ratio [OR]=0.21, 95% CI 0.02-2.33). There were no significant differences between the study arms with respect to the secondary outcomes.

Conclusions: Web-based self-help interventions supplemented by brief chat counseling are an effective alternative to face-to-face treatment and can reach a group of cannabis users who differ in their use and sociodemographic characteristics from those who enter outpatient addiction treatment.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 59948178; http://www.isrctn.com/ISRCTN59948178 (Archived by WebCite at http://www.webcitation.org/6bt01gfIr)

(J Med Internet Res 2015;17(10):e232) doi:10.2196/jmir.4860

KEYWORDS

cannabis; Internet; chat; Web based; self-help; cognitive behavioral therapy; motivational interviewing; counseling; self-control; behavioral self-management

Introduction

Web-based self-help programs that aim to reduce cannabis use might help to reach cannabis users who do not want to enter available outpatient addiction counseling services due to their fear of being stigmatized or their need to distance themselves socially from drug counselors [1]. Moreover, the limited opening hours of many outpatient addiction services might act as a barrier to care for some users [1]. It has been estimated that approximately 22% of Europeans between 15 and 64 years of age have tried cannabis. A total of 6.8% of Europeans report using cannabis in the preceding month and an estimated 3 million report daily cannabis use [2]. Switzerland has the third-highest national prevalence of cannabis use in Europe; the 12-month prevalence rate is 5.7% (men 7.8%, women 3.7%) and the 30-day prevalence rate is 2.7% (men 3.7%, women 1.7%) [3]. The age group with the highest prevalence is between 15 and 24 years of age; this group has a 12-month prevalence rate of 19.9%, and nearly one in five members from this group uses cannabis daily [3]. Daily cannabis use is associated with greater risks of developing cannabis dependence, poor mental and physical health, lower educational achievement, and decreased cognitive functioning [4]. The risks of cannabis dependence [5] and problems with cannabis use [6] are considerably higher in cannabis users with early rather than late onset of use.

Treatment demand statistics from Swiss in- and outpatient addiction treatment centers demonstrated a linear increase—from 2006 (9.9%) to 2012 (14.7%)—in new treatment entry cases for whom cannabis was the main problem substance [7]. The main group seeking treatment for cannabis use disorder mainly consists of adolescents and young adults between the ages of 15 and 24 years old (71.6%) and are predominantly male (82.5%) [8]. In Europe, cannabis is the main problem substance for almost 40% of all individuals entering addiction treatment for the first time and has been a more frequent problem than opioids since 2006 [9]. It has been estimated that about 50% of problematic cannabis users will develop cannabis dependence [5] and many of these exhibit mental health problems; however, most of them are not yet in treatment. Raising awareness of cannabis-related risks to physical health might also encourage users to reduce or quit cannabis use [10]. In general, the principle of stepped care (ie, noninvasive, low-cost interventions in which therapeutic intensity can be enhanced according to need) appears to be an appropriate means for problematic cannabis users to lower their ever-increasing health care costs [11], and this consideration is of interest in Switzerland and other industrialized countries suffering from exorbitant health costs.

An initial meta-analysis included diverse studies that mainly investigated computer- and some Web-based interventions to reduce cannabis consumption and found a small overall effect size (g=0.16, 95% CI 0.09-0.22, *P*<.001) at posttreatment. There have now been three studies on the efficacy of Web-based interventions to reduce cannabis use in problematic users. First, the German *Quit the Shit* program [12] is based on principles of self-regulation and self-control and is a solution-focused approach. This program is structured into weekly personalized feedback sessions based on participants' consumption diary entries, and intake and termination chats; the total allowed program time is 50 days. Tossmann et al [12] recruited a total of 1292 cannabis users and found significant reductions in cannabis use in their intention-to-treat (ITT) analyses, but with high attrition rates. Second, a distinct version of the program was developed that consisted of one comprehensive chat session with motivational interviewing (MI) [13] in the intervention group (n=33) versus a technical information chat in the control group (n=34). No significant differences in cannabis use were found between the study groups [14]. Third, the Australian program, *Reduce Your Use: How to Break the Cannabis Habit* [15], is a fully automated self-help intervention consisting of six modules that aim to reduce the symptoms of cannabis use disorders and which is based on cognitive behavioral therapy (CBT) [16,17], MI [13], and behavioral self-management (BSM) [18]. Its efficacy was tested in a randomized controlled trial (RCT) and compared to a psychoeducative control condition that also consisted of six modules (n=225). The frequency of cannabis use and the quantity of cannabis consumed were both reduced to a greater extent in the intervention group than in the control group at 6 weeks and at the 3-month follow-up. They achieved considerably higher participation rates at the 3-month follow-up than the German *Quit the Shit* program (54% in the intervention and 52% in the control condition) [12].

The combination of a fully automated self-help intervention based on the approaches of Rooke et al [15], together with additional individual chat sessions to reduce cannabis use, could potentially increase the efficacy of interventions for problematic cannabis users—in the sense that the use is harmful to the user

or others—as has been demonstrated for the reduction of alcohol use in problematic alcohol users [19].

Thus, the current study aims to investigate and compare the efficacy of Web-based self-help interventions—in combination with or without tailored chat counseling based on CBT, MI, and BSM—in reducing cannabis use in problematic cannabis users.

Methods

Participants

Study participants were recruited by a press release, several websites from local outpatient treatment centers, and from nightlife prevention websites that were linked to the Can Reduce website [20]. In addition, advertisements were placed in Internet forums and recruitment flyers were distributed to Swiss addiction service centers and practitioners in the Canton of Zurich. Moreover, two major Swiss commuter newspapers and one Swiss weekend newspaper published extensive reports on the Can Reduce interventions in their print media and websites. The collaboration of the Swiss Research Institute for Public Health and Addiction (ISGF) and the Arud Centers for Addiction Medicine (ARUD) as the responsible study institutions was clearly stated in all recruitment channels.

Study inclusion and exclusion criteria are depicted in Table 1. In addition to the email addresses in the registration process, participants were asked to provide their telephone numbers in case they could not be reached online for the 3-month follow-up [1]. The participant information and informed consent page from the Can Reduce website is provided in Multimedia Appendix 1.

Table 1. Inclusion and exclusion criteria and rationales.

Participant criteria	Rationales
Inclusion criteria	
Minimum age of 18 years	To ensure a minimal age of participation
Read and understand German	To ensure understanding of interventions
Internet access and a valid email address	To ensure participation
Using cannabis at least once a week over the 30 days prior to study entry	To include at least occasional users
Exclusion criteria	
Current serious psychiatric disorders or history of psychosis, schizophrenia, bipolar type I disorder, or significant current suicidal or homicidal thoughts	To avoid exacerbation of serious symptoms of these severe psychiatric disorders
Other pharmacological or psychosocial treatments for cannabis use disorders	To avoid confounding treatment effects
For women: pregnancy and breastfeeding	To avoid serious complications resulting, for example, from withdrawal symptoms

Preparatory Work

The Web-based self-help intervention, Can Reduce, was based on classical CBT approaches for treating cannabis dependence [17], MI approaches [13], and BSM [18]. A detailed description of the intervention can be found in the study protocol by Schaub et al [1]. This randomized controlled trial was registered with the International Standard Randomized Controlled Trial Number (ISRCTN) registry (ISRCTN59948178).

Can Reduce is the first self-help intervention for problematic cannabis users in Switzerland. It was developed by the authors of this publication from the ISGF and the ARUD. Both institutions are located in the Canton of Zurich, Switzerland. Study participation was free of charge. The self-help part of Can Reduce was developed according to the experiences of an earlier study in problematic cocaine users [21,22] and the Global Drug Survey cannabis meter [23], and was piloted for acceptability and usability. The piloting was organized into two steps. In the first step, we piloted Can Reduce with cannabis-using students from the University of Zurich. In the second step, we combined this with additional chat sessions with two trained psychiatrists from ARUD and four of their problematic cannabis-using patients. This pilot phase resulted in some minor changes in the interventions.

Ethical Review and Informed Consent

The protocol of the RCT was approved by the Ethics Committee of the Canton of Zurich (KEK-StV-Nr. 15/13) and was carried out in compliance with the Helsinki Declaration. Before giving informed consent, participants were informed of the following: (1) the rationale of the study, (2) study inclusion and exclusion criteria (see Table 1), (3) the three different arms and their 1:3 chance of being allocated to one of the arms, (4) the potential risks of participation, (5) safety arrangements during and after the study phase [19], (6) the inability of Can Reduce (with or without chat counseling) to replace face-to-face therapy for problematic cannabis use/abuse, (7) the circumstances under which they should contact their general practitioner or a professional from a medical advisory group; an emergency list that would be accessible at all times via an instant help button was provided as well, (8) the approval of the study by the Ethics Committee of the Canton of Zurich and their declaration of no objection (nihil obstat), and (9) their right to withdraw from the study at any time without consequences except for the loss of further compensation. Informed consent was accepted when participants clicked on all consent fields of the informed consent page and submitted the consent by clicking the submission button (see Multimedia Appendix 1).

Study Arms and Contents

There were three different study arms. The first consisted of the Web-based self-help intervention, Can Reduce, in combination with up to two individual chat counseling sessions based on MI and CBT approaches that considered the data the participants entered into the self-help intervention and individual requests. The second study arm consisted of the same intervention but without chat counseling. Study arms 1 and 2 received weekly automated motivational emails to remind the user to log in and fill out the consumption diary. Study arm 3 consisted of a classical waiting list and people in this arm received access to the self-help intervention after 3 months.

The following modules, organized into three main parts, were offered as a Web-based self-help intervention (study arms 1 and 2) and—as long as the participant did not feel an urgent need to skip to a specific module—it was recommended that they should be worked through in the order shown in Textbox 1 within the planned 6 weeks of intervention.

Textbox 1. Modules for the Can Reduce Web-based self-help intervention.

Part 1: Introduction

- Registration process
- Explanation of the "standard cannabis joint" concept and choice of the personal standard cannabis joint (see Figure 1), the cannabis consumption diary, and the automated reminder emails
- Examination of the pros and cons resulting from a change in cannabis consumption patterns and further principles of motivational interviewing to address motivation, followed by setting an appropriate target value for overall cannabis use, which is to be reached by the end of the intervention
- Explanation of the *My Can Reduce* folder
- Explanation of the emergency button for immediate responses to frequently asked questions and access to emergency contacts

Part 2: Key Modules (participants are encouraged to complete these modules in the order presented below; see Figure 2)

- Module 1: Strategies for goal achievement
- Module 2: Identifying risk situations
- Module 3: Dealing with cannabis craving
- Module 4: Dealing with relapses

Part 3: Further Modules (participants are encouraged to complete at least two, in any order)

- Module 5: Tobacco smoking during the reduction in cannabis use
- Module 6: Saying "no" to foster refusal skills
- Module 7: Dealing with burdens
- Module 8: Preserving achievements

Figures 1 and 2 show screenshots of the Can Reduce Web-based intervention website. The following were also provided: a glossary that explained the terms, definitions, and concepts used in the intervention; a knowledge base about the history of cannabis use; the effects and risks of cannabis use; concurrent mental health problems; and the enhanced risks when cannabis is mixed with tobacco and smoked, as in a previously developed and positively evaluated cannabis group smoking cessation program [10,24] (study arm 1 and 2). The knowledge base also included harm reduction techniques with recommendations for the use of cannabis [25,26].

The additional (up to two) chat counseling sessions with a scheduled duration of 20 to 30 minutes in study arm 1 supported behavioral change according to MI, discussed the modules of the Web-based self-help part based on MI and CBT, and reviewed the development of the consumption diary. Invitations to chat sessions were sent by the counselors according to a predefined procedure between weeks 1 and 2 for the first and between weeks 4 and 6 for the second chat session. The chats took place within the website in a small box at the bottom right corner, while keeping the content of the webpage in view (see Figure 2). It was initially planned that the structure of these chat sessions should be fixed [1]. However, as a result of the counselor supervision sessions, the structure of the chat session was made more flexible and more dependent on the participants' needs and served as a checklist for the counselors in order to ensure that they covered all of the relevant contents.

The chat counselors received quarterly supervision sessions and consisted of trained MI counselors, mainly psychologists or psychiatrists with advanced or completed further education, with at least one year of experience in treating cannabis-abusing patients face to face. Specific quality standards were developed for addiction chat counseling and implemented for this study in the chat counselor supervision based on the study on the development of a European Union framework for minimum quality standards and benchmarks in drug demand reduction treatment quality standards [27] and the Swiss national addiction counseling quality standards [28].

To optimize and manage their interactions with clients, counselors had access to a specific user management area to add arranged chat dates, define statuses, and add personal comments about their clients. With this tool, counselors could

follow their clients' progress in reducing their cannabis use through clearly arranged charts, and look up previous chat histories. Specific lists helped counselors track their clients (eg, a list with *all users*, *my clients*, or *my upcoming chat sessions*).

The Web-based self-help intervention and the subsequent tailored chat counseling aimed to reduce cannabis use. However, those participants who sought cannabis abstinence were also encouraged to make step-by-step reductions until full abstinence was reached. In accordance with the counselor supervision group, we deviated from the study protocol [1] by introducing the option to dispense with a second chat session if a participant and his/her counselor agreed that another chat would not be needed.

Participants randomized to the waiting list had the opportunity to participate in the Web-based self-help intervention 3 months after registration.

Figure 1. Screenshot of the Can Reduce Web-based intervention, showing the decision on the standard cannabis joint prior to the first consumption diary entry.

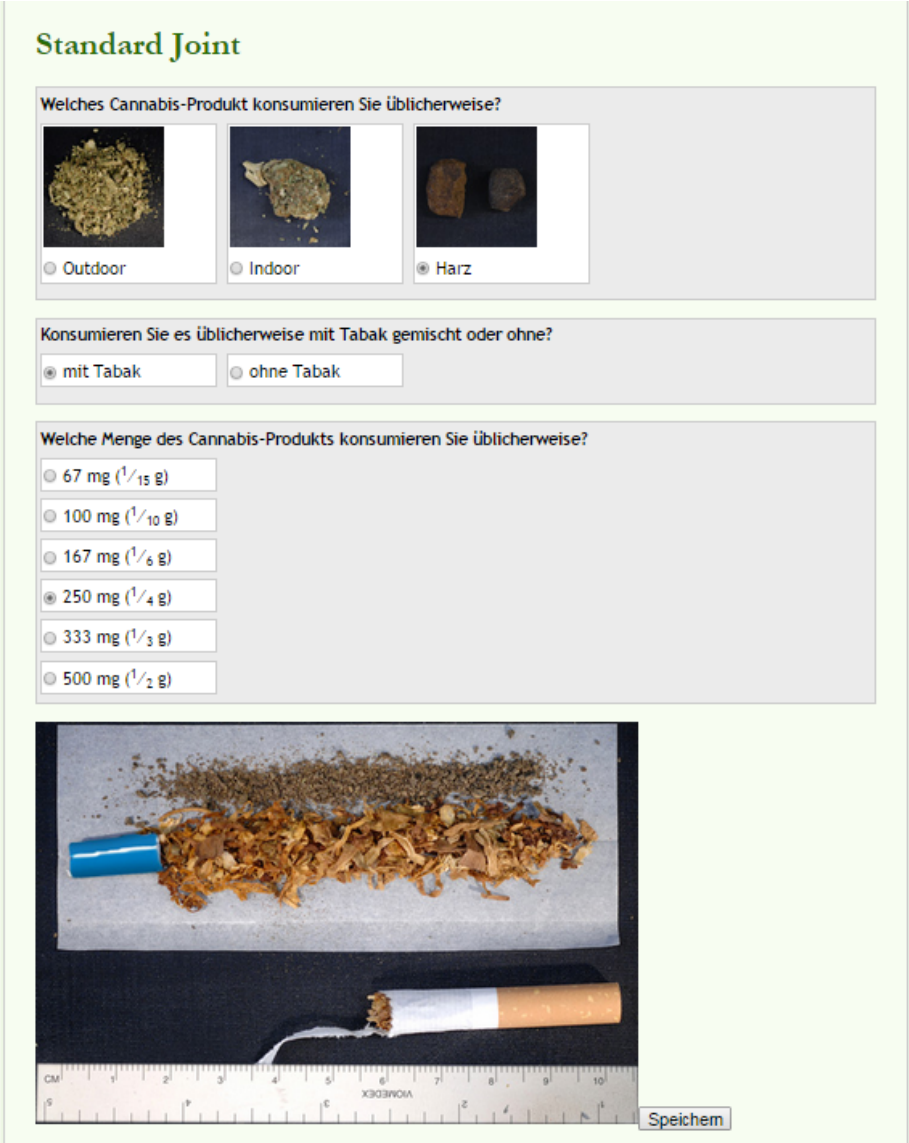
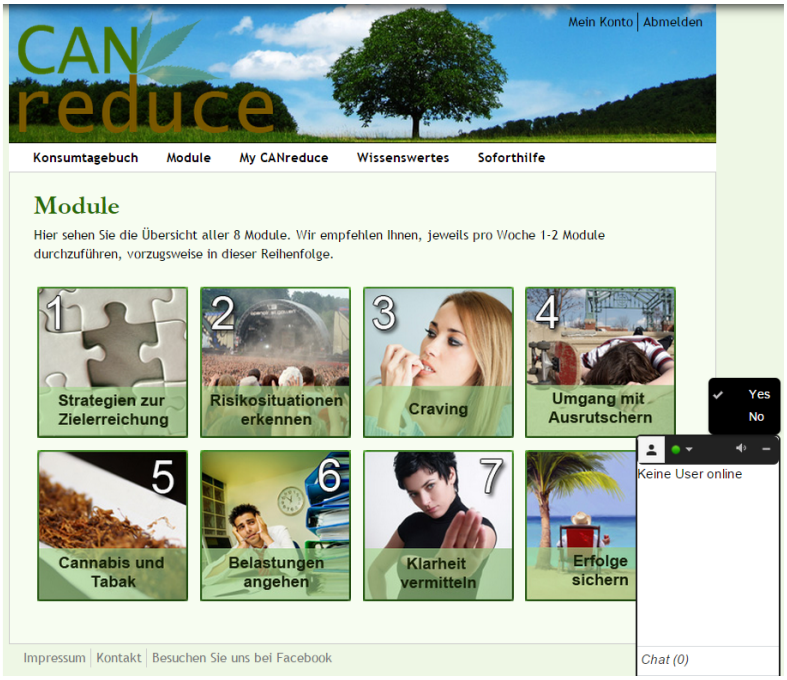


Figure 2. Main menu of the Can Reduce Web-based intervention's study arm 1 with self-help plus chat counseling that took place within the website in a small box at the bottom right corner.



Detailed Study Hypotheses

This study aimed at comparing the efficacy of a Web-based self-help intervention alone or combined with chat counseling in the reduction of the cannabis use of problematic cannabis users within a three-arm randomized controlled trial with assessments at baseline and 3-month follow-up (see Multimedia Appendix 2 for the CONSORT-EHEALTH checklist [29]).

We hypothesized that Web-based interventions—which are more interactive—would be more effective than less interactive interventions in reducing cannabis use among problematic cannabis users. We tested the following detailed study hypotheses with respect to the main outcome (ie, the reduction of the weekly cannabis used between the baseline and the 3-month follow-up):

1. Tailored chat-based counseling in combination with Web-based self-help for the reduction of cannabis use (study arm 1) is more effective than the waiting list control condition (study arm 3).
2. Web-based self-help for the reduction of cannabis use (study arm 2) is more effective than the waiting list control condition (study arm 3).
3. Chat-based counseling in addition to Web-based self-help for the reduction of cannabis use (study arm 1) exhibits a trend to be more effective than Web-based self-help alone (study arm 2).

Measurement Instruments

The primary outcome measure was the recorded quantity of cannabis use in the previous 7 days, quantified in individually standardized cannabis joint sizes, and as specified in the consumption diary (see Table 2 and Schaub et al for further details [1]). In the first step, participants chose between three different cannabis forms presented in photographs—low-potency cannabis plant, high-potency cannabis plant, or cannabis resin (see Figure 1). In the second step, five different standard joints for each category were presented (1/10 g, 1/6 g, 1/4 g, 1/3 g, 1/2 g; pictures came from the Global Drug Survey cannabis meter [23]); these joints were either pure cannabis or cannabis mixed with tobacco. A standard tobacco cigarette, a ruler with centimeter and millimeter scales, the fraction amount in grams, and an open 10 cm paper prepared to roll a joint and containing the cannabis plant-/resin-tobacco mixture or pure cannabis were presented. Participants chose which picture most closely approximated the cannabis joints they most often smoke. The chosen picture was placed in the individual consumption diary (see Figure 1), and participants were asked to convert the quantities of cannabis they smoked into units relative to that picture if they exceptionally consumed cannabis in forms other than their common standard joint. As this kind of outcome assessment has not previously been used in an efficacy trial, we also considered the number of cannabis use days in the last 7 days as a primary outcome [1].

The following secondary outcome instruments were applied:

1. The Cannabis Use Disorders Identification Test (CUDIT), which is a 10-item questionnaire [30] that was constructed by adapting the Alcohol Use Disorders Identification Test [31]. To cover the length of the trial, this instrument was adapted to focus on the last 3 months in its planned assessments (baseline and 3-month follow-up).

2. The Severity of Dependence Scale (SDS), which is a five-item questionnaire that measures the severity of cannabis dependence. Each of the five items is scored on a 4-point scale (0-3). The total score is obtained by adding the ratings on all five items. High scores indicate high levels of dependency [32].

3. The Cannabis Withdrawal Scale (CWS) [33], which is a 19-item questionnaire containing statements that describe cannabis withdrawal symptoms within the last 24 hours on an 11-point scale (0-10).

4. The Cannabis Craving Symptoms questionnaire (CCS-7), which is a seven-item questionnaire [34] derived from the Marijuana Craving Questionnaire [35]. Each item is rated on a 7-point scale (1-7).

5. The Fragebogen Substanzanamnese (FDA), which is a questionnaire that ascertains the number of years of consumption over the lifetime, the past month's consumption, and the manner of consumption for the Diagnostic and Statistical Manual of Mental Disorders' substances of abuse. This measure was derived from the Europe Addiction Severity Index [36].

6. The short version of the Mental Health Inventory (MHI-5) [37], which is a validated and user-friendly self-assessment questionnaire that assesses recent mental distress and self-reported diagnoses of depression.

None of the secondary outcome instruments has yet been specifically validated for Internet use. Intervention satisfaction for all modules, the diary, the chat, the knowledge base, the instant help, and the overall satisfaction was ascertained on a 4-point scale, ranging from *not at all useful* to *very useful*. Finally, intervention participation was assessed for completed modules each time a participant pressed the *back to the main menu* button at the very end of a module. Retention was calculated as the percentage of days per week a user entered any number of cannabis use in the diary.

Table 2. Study measurements and instruments.

Assessments/instruments	Baseline	1 week	3 weeks	6 weeks	3-month follow-up
Sociodemographics	x				
MHI-5[a]	x				x
Quantity of cannabis use[b]	x	x	x	x	x
Frequency of cannabis use[b]	x	x	x	x	x
CUDIT[c]	x				x
SDS[d]	x			x	x
FDA[e]	x			x	x

[a]Mental Health Inventory (MHI-5).

[b]7-day point prevalence values of the quantity (in common standard joints) and frequency (the number of days on which cannabis is used) of cannabis use were derived from the consumption diary for the preceding 7 days.

[c]Cannabis Use Disorders Identification Test (CUDIT).

[d]Severity of Dependence Scale (SDS).

[e]Fragebogen Substanzanamnese (FDA).

Sample Size

Based on results of the study of Rooke et al [15], we expected small to medium effect sizes of at least 0.30 (Cohen's d) for the reduction in the quantity of cannabis used and the frequency of cannabis use between study arm 2 (Web-based self-help without chat counseling) and study arm 3 (waiting list control) between baseline and follow-up assessment, and greater effects between study arms 1 and 3. We estimated a sample size of 89 in each study group that would have 80% power (*F* test, alpha = 5%) to detect these differences, as based on calculations with G*Power software version 3.1. Therefore, we aimed to recruit a total of 267 participants [1]. We had no reference values for the expected differences in effects between study arms 1 and 2 and thus planned an exploratory study of effect sizes in case we failed to reach significance for these study arm comparisons.

Randomization and Allocation

Once participants had completed their baseline assessment, they were randomized by a computer program in a 1:1:1 ratio to one of three parallel groups. As the participant information offered full transparency on the three study arms in our nonblinded design, we anticipated a risk that some participants might register another account, in an effort to change their assignment and access a different study arm. In that case, the participant remained in the initially assigned study arm for the rest of the day, as based on his or her IP address.

Statistical Methods

Data were analyzed according to the intention-to-treat principle. For the ITT analyses, in departure from the study protocol, we applied multiple imputation procedures of R (R Foundation for Statistical Computing, Vienna, Austria) in Amelia II that have been demonstrated to outperform other imputation methods [38]. For each study arm, we performed 50 separate imputations using the following as imputation variables: sex, age, education, origin, years of cannabis use, number of finished modules, the baseline variables for frequency and quantity of cannabis use, alcohol use in the last 30 days (risky and normal), SDS, CUDIT, and MHI-5. Baseline measurements were compared between the three study arms and study participants were compared with people entering addiction treatment according to data from the Swiss *addiction, care and therapy information* (act-info) monitoring statistics. Depending on the scale of the corresponding outcome, Mann-Whitney U tests, chi-square tests, or analyses of variance (ANOVA) were calculated via SPSS version 22.0 (IBM Corporation). The calculation of the changes between baseline and the 3-month follow-up was modified from the protocol, as there were a considerable number of missing values at the 6-week assessment. Regression analyses

in R were used for the calculated differences between 3-month follow-up and baseline, using the corresponding baseline variables as control variables. Results from the imputed dataset were cross-checked with the nonimputed dataset in the latter analyses. In departure from the study protocol, we dispensed with analyzing the Cannabis Withdrawal Scale and the Cannabis Craving Symptoms questionnaire data [1], as very low numbers of questionnaires were completed at intervention weeks 3 and 6. In the study dropout analysis, we conducted regression analyses to investigate the interaction effect of relevant baseline characteristics (ie, sociodemographic and consumption characteristics) between those who did and those who did not provide a 3-month follow-up. These analyses were conducted for the total sample and for each study arm separately. Similar analyses were conducted in the subgroup analyses.

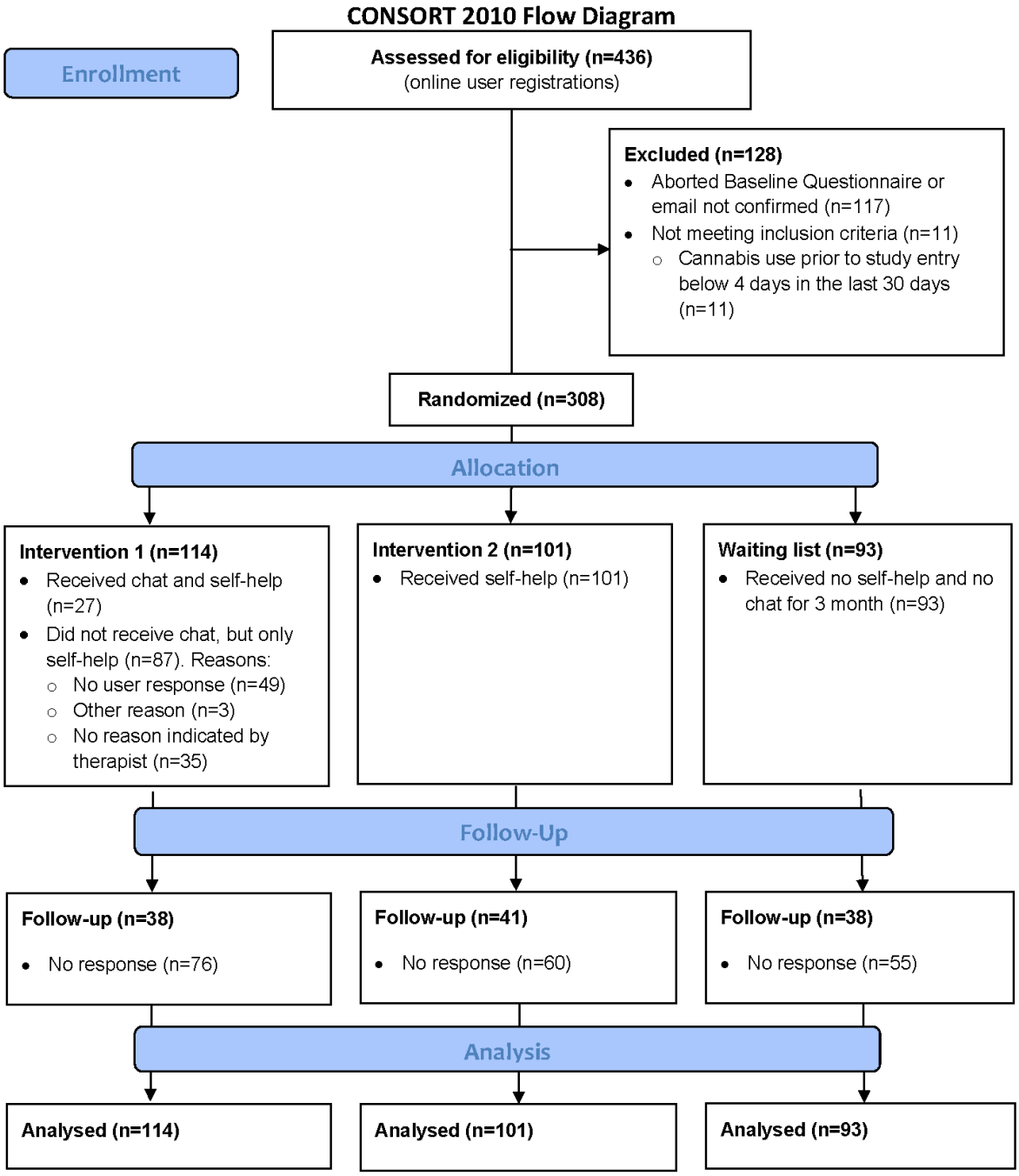
Results

Participant Flow

Figure 3 provides an overview of the trial flow. Recruitment started in the beginning of June 2014 and ended on February 28, 2015, after exceeding the total estimated number of 267 participants. Of the 436 Can Reduce registrants recruited, 308 (70.6%) were allocated to one of the three study arms.

Three months after the baseline assessment, participants were invited by email to log in and complete the final study assessment; they were reimbursed with €40 (via an online voucher or an online charitable donation). The follow-up assessment was performed in three steps. First, participants were invited via email to participate in the assessment. Up to three reminders were sent. Those participants who failed to complete the 3-month follow-up despite these reminders were contacted via telephone and offered an interview by study collaborators. Those participants who refused a telephone interview were offered an interview on the primary outcome only. Finally, 117 out of 308 participants (38.0%) could be followed up with.

Figure 3. CONSORT-EHEALTH trial flowchart: overview of the participant flow for this trial.



Participants' Baseline Characteristics

Table 3 provides an overview of the participants' characteristics and comparisons between the three study arms at baseline assessment. In comparison with participants whose main problem substance was cannabis in the Swiss treatment monitoring statistics (act-info) in 2013 [7], Can Reduce participants demonstrated a similar gender distribution (75.3% males in Can Reduce vs 82.4% act-info, U=1.342, *P*=.18), tended to be older within the age groups between 20 and 69 years old (U=2.296, *P*=.02), and reported a higher number of cannabis use days in the 7 days prior to intervention start (70.9% daily use in Can Reduce vs 41.4% act-info; 20.2% 4-6 days per week Can Reduce vs 10.5% act-info; 5.2% 2-3 days per week Can Reduce vs 21.7% act-info; 3.8% 1 day a week Can Reduce vs 26.3% act-info; chi-square [df 2]=4.0, *P*=.046).

Table 3. Baseline characteristics of participants.

Characteristics	Study arm 1[a] (n=114)	Study arm 2[b] (n=101)	Study arm 3[c] (n=93)	Total (n=308)	χ^2, ANOVA[d], or Kruskal-Wallis test	*P*
Sex, n (%)					χ^2_2=4.3 (n=308)	.12
Female	35 (30.7)	24 (23.8)	17 (18)	76 (24.7)		
Male	79 (69.3)	77 (76.2)	76 (82)	232 (75.3)		
Age in years, mean (SD)	28.4 (9.6)	30.2 (9.2)	31.0 (11.1)	29.8 (10.0)	$F_{2,308}$=1.940	.15
Age range, n (%)					χ^2_2=3.9 (n=308)	.14
≤20 years	24 (21.1)	12 (11.9)	18 (19)	54 (17.5)		
21-25 years	31 (27.2)	19 (18.8)	13 (14)	63 (20.5)		
26-30 years	16 (14.0)	29 (28.7)	19 (20)	64 (20.8)		
31-35 years	17 (14.9)	18 (17.8)	15 (16)	50 (16.2)		
36-40 years	14 (12.3)	10 (9.9)	11 (12)	35 (11.4)		
41-45 years	6 (5.3)	5 (5.0)	7 (8)	18 (5.8)		
46+ years	6 (5.3)	8 (7.9)	10 (11)	24 (7.8)		
Highest education, n (%)					χ^2_{10}=8.6 (n=308)	.57
Not specified	4 (3.5)	3 (3.0)	5 (5)	12 (3.9)		
Primary school	18 (15.8)	12 (11.9)	11 (12)	41 (13.3)		
Apprenticeship	43 (37.7)	38 (37.6)	41 (44)	122 (39.6)		
Secondary school	19 (16.7)	13 (12.9)	17 (18)	49 (15.9)		
Technical college	18 (15.8)	26 (25.7)	13 (14)	57 (18.5)		
University	12 (10.5)	9 (8.9)	6 (7)	27 (8.8)		
Origin, n (%)					χ^2_6=8.1 (n=308)	.23
Canton of Zurich	52 (45.6)	33 (32.7)	42 (45)	127 (41.2)		
Other cantons	53 (46.5)	61 (60.4)	48 (52)	162 (52.6)		
Germany	8 (7.0)	5 (5.0)	3 (3)	16 (5.2)		
Other countries	1 (0.9)	2 (2.0)	0 (0)	3 (1.0)		
CUDIT[e], mean (SD)	19.8 (5.8)	19.7 (6.4)	19.1 (6.2)	19.6 (6.1)	$F_{2,308}$=0.37	.69
SDS[f], mean (SD)	7.7 (3.5)	7.5 (3.6)	7.3 (3.2)	7.5 (3.4)	$F_{2,308}$=0.37	.69
MHI-5[g], mean (SD)	54.0 (19.3)	53.9 (20.0)	55.1 (22.6)	54.3 (20.5)	$F_{2,308}$=0.11	.90
Number of years of substance use, mean (SD)						
Cannabinoids	9.6 (7.4)	10.9 (7.6)	12.6 (10.0)	10.9 (8.4)	$F_{2,305}$=3.29	.04[h]
Risky alcohol use[i]	2.5 (5.6)	2.6 (5.3)	2.7 (6.4)	2.6 (5.7)	$F_{2,228}$=0.03	.97
Cocaine	1.1 (4.3)	1.4 (3.8)	0.8 (1.8)	1.1 (3.4)	$F_{2,222}$=0.67	.52
Amphetamines	0.7 (2.0)	1.1 (3.1)	0.6 (2.0)	0.8 (2.4)	$F_{2,208}$=0.91	.40
Substance use in the last 30 days, n (%)						
Cannabinoids	112 (98.2)	100 (99.0)	93 (100)	305 (99.0)	Not computable (no variance)	N/A[j]
Risky alcohol use[i]	40 (35.1)	26 (25.7)	31 (33)	97 (31.5)	χ^2_2=2.6 (n=226)	.28
Tranquilizers	7 (6.1)	8 (7.9)	5 (5)	20 (6.5)	χ^2_2=0.6 (n=215)	.74

Characteristics	Study arm 1[a] (n=114)	Study arm 2[b] (n=101)	Study arm 3[c] (n=93)	Total (n=308)	χ^2, ANOVA[d], or Kruskal-Wallis test	*P*
Cocaine	7 (6.1)	14 (13.9)	10 (11)	31 (10.1)	χ^2_2=2.7 (n=223)	.26
Amphetamines	16 (14.0)	13 (12.9)	14 (15)	43 (14.0)	χ^2_2=0.2 (n=221)	.90
Hallucinogens	6 (5.3)	4 (4.0)	4 (4)	14 (4.5)	χ^2_2=0.6 (n=210)	.76
Heroin	0 (0)	1 (1.0)	0 (0)	1 (0.3)	χ^2_2=1.8 (n=201)	.40
Methadone	1 (0.9)	1 (1.0)	3 (3)	5 (1.6)	χ^2_2=1.9 (n=197)	.40
Others	4 (3.5)	1 (1.0)	1 (1)	6 (1.9)	χ^2_2=3.1 (n=198)	.21

[a]Self-help with chat.

[b]Self-help without chat.

[c]Waiting list.

[d]Analysis of variance (ANOVA).

[e]Cannabis Use Disorders Identification Test (CUDIT) scores range from 0 to 40 with a cutoff of >8 for a cannabis use disorder.

[f]Severity of Dependence Scale (SDS) scores range from 0 to 15 with a cutoff of ≥4 for cannabis dependence.

[g]Mental Health Inventory (MHI-5): higher values represent improved symptoms. MHI-5 values range from 0 to 100 with a cutoff of <70 for clinically relevant symptoms.

[h]*P*<.05, represents a significant value.

[i]Risky alcohol use was defined as five or more standard drinks per day on at least three days per week. A standard drink was defined as 5 cl spirits, 15-20 cl wine, or 33-45 cl beer.

[j]Not applicable (N/A).

Intervention Participation and Retention

Figure 4 depicts the module completion by participants in study arms 1 and 2. Participants in the self-help with chat study arm completed a mean of 3.2 modules and 27 out of 114 (23.7%) of the participants received at least one chat session. Participants in the self-help without chat study arm completed similar numbers of self-help modules (U=-1.189, *P*=.23). Participants in study arm 1 more frequently completed the consumption diary than those in study arm 2 during their recommended 6 intervention weeks (U=-2.375, *P*=.02; see Figure 5). Of the 27 users in study arm 1 who received chat counseling sessions, 23 (85%) received one session and 4 (15%) received two sessions.

Figure 4. Module completion rate for study arms 1 (self-help with chat) and 2 (self-help without chat).

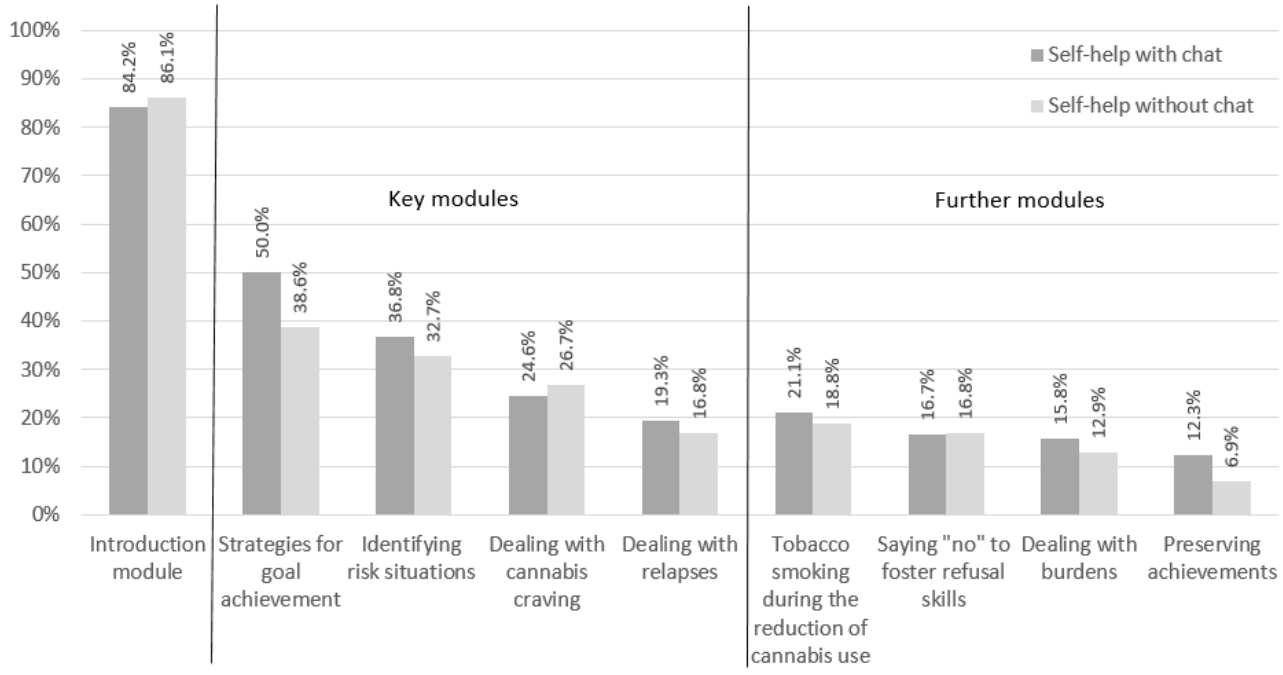
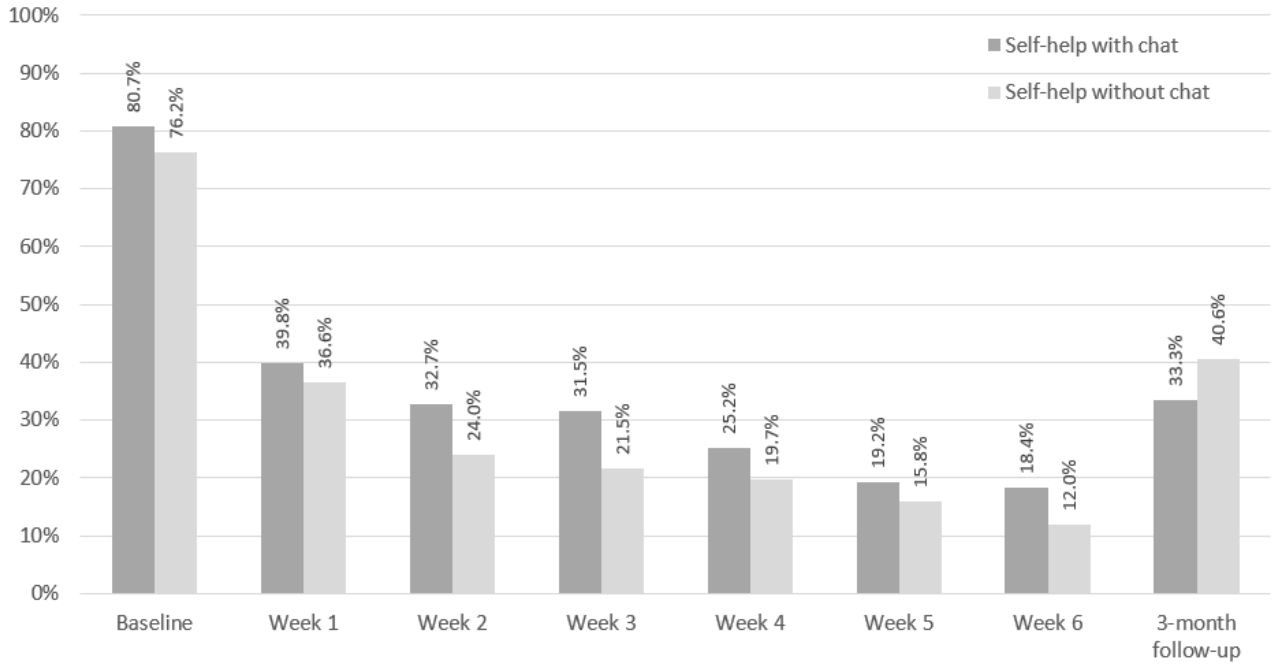


Figure 5. Study retention based on the weekly completion of the consumption diary for study arms 1 (self-help with chat) and 2 (self-help without chat) between baseline and week 6, including 3-month follow-up completion rate.



Main Outcomes

Figure 6 depicts the mean numbers of cannabis use days per week and Figure 7 the mean weekly quantity of cannabis used in standard joints according to the consumption diary, between baseline and follow-up for all three study arms and based on the nonimputed dataset.

The differences in cannabis use between baseline and the 3-month follow-up, as expressed by the mean number of cannabis use days per week and based on the imputed data, differed between self-help without chat versus self-help with chat (beta= -0.75, SE = 0.32, *t*=-2.39, *P*=.02, d=0.34, 95% CI 0.07-0.61), and between self-help with chat versus waiting list (beta= 0.70, SE = 0.32, *t*=2.16, *P*=.03, d=0.20, 95% CI -0.07 to 0.47), but not between self-help without chat versus waiting list (beta= -0.05, SE = 0.33, *t*=-0.16, *P*=.87, d=-0.14, 95% CI -0.43 to 0.14). In contrast, we only observed one trend to a significant difference in the weekly quantity of standard joints in the comparison of self-help with chat versus waiting list in the imputed dataset (beta = 4.73, SE = 2.50, *t*=1.89, *P*=.06, d=0.09, 95% CI -0.19 to 0.36; see Tables 4 and 5).

Figure 6. Cannabis use days per week according to the consumption diary between baseline and 3-month follow-up for all three study arms based on the nonimputed dataset.

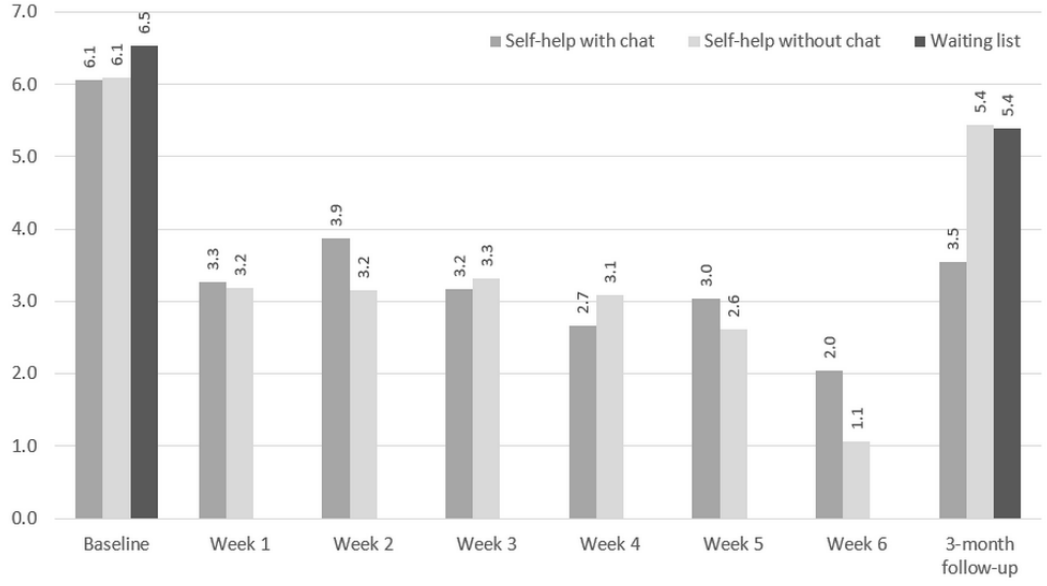
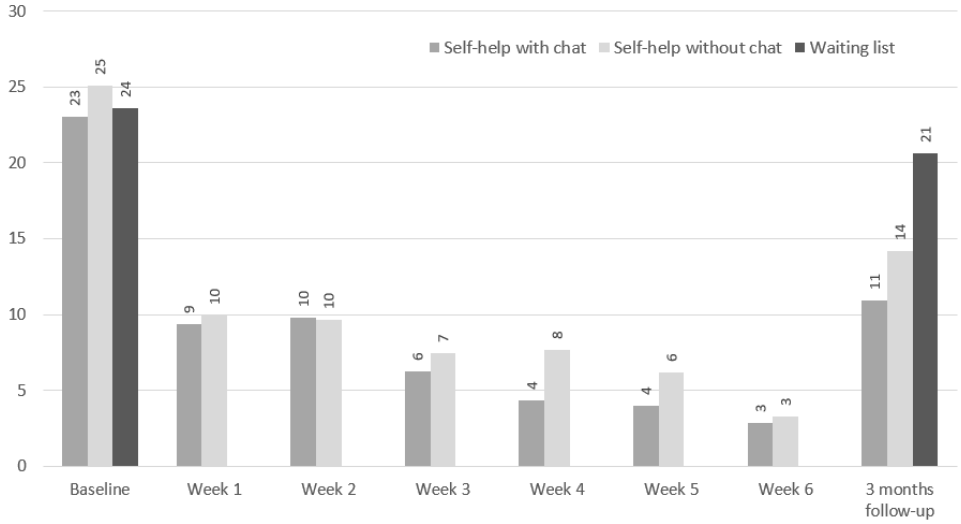


Figure 7. Weekly quantity of cannabis used in number of standardized cannabis joints between baseline and 3-month follow-up for all three study arms based on the nonimputed dataset.



Secondary Outcomes

There were no significant differences in the group comparisons in the secondary outcomes (see Tables 4 and 5). We observed slight improvements in mental health (MHI-5), cannabis use disorders (CUDIT), and severity of dependence (SDS) in all three groups (see Table 4; pre/post comparisons not reported). Assessment of the intervention satisfaction was completed by only a few participants at 6 months past baseline and we therefore omit group comparisons. Not surprisingly, those who remained in the active study arms rated their satisfaction as high (eg, intervention satisfaction in general [19/308, 6.2%]: very satisfied 42% [8/19], quite satisfied with most of the intervention 42% [8/19], quite unsatisfied with most of the intervention 11% [2/19], quite unsatisfied 5% [1/19]).

Table 4. Number of participants and mean and standard deviation changes from the imputed (50 imputations) and complete case datasets between baseline and 3-month follow-up.

Outcomes	Study arm 1 (self-help with chat) (n=114), mean (SD)		Study arm 2 (self-help without chat) (n=101), mean (SD)		Study arm 3 (waiting list) (n=93), mean (SD)	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Frequency of cannabis use [a]						
Imputed data	6.0 (1.6)	4.6 (2.1)	6.0 (1.6)	5.3 (1.8)	6.3 (1.0)	5.3 (1.8)
Complete cases	6.1 (1.6)	3.8 (3.0)	6.1 (1.7)	5.5 (2.3)	6.7 (0.9)	5.3 (2.5)
Quantity of cannabis use [b]						
Imputed data	22.3 (14.8)	13.3 (12.0)	23.1 (23.1)	14.4 (11.8)	25.8 (18.7)	18.6 (17.7)
Complete cases	23.0 (15.1)	10.9 (13.8)	25.1 (25.2)	14.2 (13.3)	23.6 (13.2)	20.7 (23.7)
CUDIT [c]						
Imputed data	19.8 (5.8)	16.6 (7.1)	19.7 (6.4)	15.6 (6.7)	19.1 (6.2)	16.6 (6.4)
Complete cases	19.8 (5.8)	12.6 (8.4)	19.7 (6.4)	13.0 (7.4)	19.1 (6.2)	16.0 (7.2)
SDS [d]						
Imputed data	7.7 (3.5)	6.3 (3.3)	7.5 (3.6)	6.2 (3.1)	7.3 (3.1)	6.3 (3.3)
Complete cases	7.7 (3.5)	5.3 (3.8)	7.5 (3.6)	6.0 (3.3)	7.3 (3.1)	5.9 (3.8)
MHI-5 [e]						
Imputed data	53.9 (19.3)	58.1 (18.2)	53.9 (20.0)	60.4 (19.1)	55.1 (22.6)	59.4 (19.4)
Complete cases	53.9 (19.3)	62.4 (19.8)	53.9 (20.0)	63.4 (20.4)	55.1 (22.6)	64.6 (18.3)
Alcohol use in the last 30 days (risky)						
Imputed data	3.4 (6.2)	2.8 (2.8)	2.4 (5.0)	2.2 (3.0)	4.5 (7.9)	3.3 (4.0)
Complete cases	3.4 (7.0)	1.6 (2.6)	2.5 (5.8)	1.0 (2.6)	4.5 (8.7)	2.1 (4.7)

[a]Based on the weekly number of cannabis use days according to the consumption diary.

[b]Based on the weekly number of standard cannabis joints according to the consumption diary.

[c]Cannabis Use Disorders Identification Test (CUDIT) scores range from 0 to 40 with a cutoff of >8 for a cannabis use disorder.

[d]Severity of Dependence Scale (SDS) scores range from 0 to 15 with a cutoff of ≥4 for cannabis dependence.

[e]Mental Health Inventory (MHI-5): higher values represent improved symptoms. MHI-5 values range from 0 to 100 with a cutoff of <70 for clinically relevant symptoms.

Table 5. Results for the between-study arm[a] comparisons from the linear (and logistic) regression models and calculated effect sizes based on the imputed dataset (50 imputations).

Characteristics	beta	SE	*t*	*P*	Cohen's d (95% CI)
Frequency of cannabis use [b]					
(Intercept)	-3.95	0.58	-6.76	<.001	
Arm 1 vs arm 3	0.70	0.32	2.16	*.03* [c]	*0.20* (-0.07 to 0.47)
Arm 2 vs arm 3	-0.05	0.33	-0.16	.87	-0.14 (-0.43 to 0.14)
(Intercept)	-3.25	0.56	-5.79	<.001	
Arm 2 vs arm 1	-0.75	0.32	-2.39	*.02*	*0.34* (0.07 to 0.61)
Quantity of cannabis use [d]					
(Intercept)	-14.50	2.24	-6.46	<.001	
Arm 1 vs arm 3	4.73	2.50	1.89	*.06*	0.09 (-0.19 to 0.36)
Arm 2 vs arm 3	3.77	2.42	1.56	.12	0.06 (-0.22 to 0.35)
(Intercept)	-9.78	1.92	-5.09	<.001	
Arm 2 vs arm 1	-0.96	2.43	-0.39	.69	0.01 (-0.26 to 0.28)
CUDIT [e]					
(Intercept)	-10.39	1.61	-6.46	<.001	
Arm 1 vs arm 3	0.24	1.29	0.19	.85	0.09 (-0.18 to 0.37)
Arm 2 vs arm 3	1.19	1.20	0.99	.32	0.21 (-0.07 to 0.49)
(Intercept)	-10.14	1.68	-6.05	<.001	
Arm 2 vs arm 1	0.95	1.16	0.82	.41	-0.12 (-0.39 to 0.14)
SDS [f]					
(Intercept)	-4.68	0.64	-7.34	<.001	
Arm 1 vs arm 3	0.03	0.58	0.05	.96	0.08 (-0.19 to 0.36)
Arm 2 vs arm 3	0.10	0.56	0.17	.86	0.07 (-0.21 to 0.35)
(Intercept)	-4.65	0.65	-7.19	<.001	
Arm 2 vs arm 1	0.07	0.55	0.13	.90	0.02 (-0.25 to 0.28)
MHI-5 [g]					
(Intercept)	-43.91	4.33	-10.15	<.001	
Arm 1 vs arm 3	0.96	3.44	0.28	.78	0.01 (-0.27 to 0.28)
Arm 2 vs arm 3	-1.38	3.42	-0.40	.69	-0.09 (-0.38 to 0.19)
(Intercept)	-42.95	4.14	-10.37	<.001	
Arm 2 vs arm 1	-2.34	3.28	-0.71	.48	0.11 (-0.16 to 0.38)
Alcohol use in the last 30 days (risky)					
(Intercept)	-2.84	0.56	-5.09	<.001	
Arm 1 vs arm 3	0.32	0.63	0.52	.61	-0.10 (-0.38 to 0.17)
Arm 2 vs arm 3	0.83	0.73	1.14	.25	-0.16 (-0.44 to 0.12)
(Intercept)	-2.52	0.46	-5.46	<.001	
Arm 2 vs arm 1	0.51	0.61	0.84	.40	0.06 (-0.20 to 0.33)

[a]Study arm 1: self-help with chat; study arm 2: self-help without chat; study arm 3: waiting list.

[b]Based on the weekly number of cannabis use days according to the consumption diary.

[c]Significant and borderline significant differences and effect sizes are in italics.

[d]Based on the weekly number of standard cannabis joints according to the consumption diary.

[e]Cannabis Use Disorders Identification Test (CUDIT) scores range from 0 to 40 with a cutoff of >8 for a cannabis use disorder.

[f]Severity of Dependence Scale (SDS) scores range from 0 to 15 with a cutoff of ≥4 for cannabis dependence.

[g]Mental Health Inventory (MHI-5): higher values represent improved symptoms. MHI-5 scores range from 0 to 100 with a cutoff of <70 for clinically relevant symptoms.

Dropout Analysis

Dropouts at follow-up did not differ from completers with respect to the following baseline variables: gender (t=1.34, P=.16), age (t=-0.24, P=.81), years of cannabis use (t=0.18, P=.86), frequency of cannabis use in the preceding 30 days (t=0.22, P=.83), the weekly number of standardized cannabis joints used (t=1.20, P=.42), the SDS (t=-1.52, P=.13), the CUDIT (t=0.49, P=.63), alcohol use in the preceding 30 days (t=1.20, P=.23), risky alcohol use in the preceding 30 days (t=1.56, P=.12), and the MHI-5 (t=0.40, P=.69).

Significantly more participants could be followed up who received at least one chat session compared to those who could not be contacted at the 3-month follow-up (17.0% vs 5.5%, chi-square [df 2]= 7.5, P=.001).

Dropouts did not differ between the three study arms with respect to gender (F_2 =0.04, P=.96), age (F_2 = 1.13, P=.27), years of cannabis use (F_2 = 0.81, P=.79), frequency of cannabis use in the preceding 30 days (F_2 = 0.91, P=.59), the standardized cannabis use quantity (F_2 = 0.93, P=.60), the SDS (F_2 = 1.20, P=.23), the CUDIT (F_2 = 0.94, P=.58), alcohol use in the preceding 30 days (F_2 = 0.57, P=.97), risky alcohol use in the preceding 30 days (F_2 = 0.48, P=.98), and the MHI-5 (F_2 = 1.00, P=.47) at baseline.

Nonintended Results

Although not intended as an outcome measure, we also offered cannabis abstention in the study protocol for those participants who wished to achieve this [1]. Self-reported 7-day point prevalence abstinence was significantly higher in self-help with chat (8.8%) than in the self-help without chat study arm (2.0%; beta = -1.56, SE = 0.79, P=.05, odds ratio [OR] = 0.21, 95% CI 0.02-2.33), but not between the self-help study arm with chat and the waiting list control group (4.3%; beta = 0.76, SE = 0.61, P=.21, OR = 2.14, 95% CI 0.86-5.30; see Tables 6 and 7).

Table 6. Number of participants in three study arms at each time point.

Study time point	Study arm 1 (self-help with chat) (n=114), n (%)	Study arm 2 (self-help without chat) (n=101), n (%)	Study arm 3 (waiting list) (n=93), n (%)
Week 1	9 (7.9)	12 (11.9)	N/A[a]
Week 6	8 (7.0)	9 (8.9)	N/A
Follow-up	9 (8.8)	2 (2.0)	4 (4)

[a]Not applicable (N/A).

Table 7. Self-reported abstinence between groups and with the corresponding logistic regression.

Abstinence at follow-up	beta	SE	t	P	OR[a] (95% CI)
(Intercept)	0.04	0.02	1.88	.06	
Arm 1 vs arm 3	0.76	0.61	1.25	.21	2.14 (0.86-5.30)
Arm 2 vs arm 3	-0.80	0.88	-0.91	.36	0.45 (0.11-1.78)
(Intercept)	-2.34	0.33	-7.07	<.001	
Arm 2 vs arm 1	-1.56	0.79	-1.98	*.05* [b]	0.21(0.02-2.33)

[a]Odds ratio (OR).

[b]Borderline significant difference is shown in italics.

Subgroup Analyses

Participants in study arm 1 who received at least one chat session exhibited lower changes in their entries in the consumption diary. This meant that they took longer to complete the consumption diary and exhibited higher retention (change in mean 0.3 vs 0.5; beta = -0.28, SE = 0.12, P=.03, 95% CI -0.66 to -0.53) than those who did not receive the chat session in study arm 1. In line with this, they completed twice as many modules (mean 5.4, SD 2.8 vs mean 2.5, SD 2.1; t=5.45, df = 96, P<.001, 95% CI 1.88-4.00). Regarding cannabis use, these two subgroups did not differ in their reduction in frequency (change in mean 3.3 vs 1.9; beta = -1.38, SE = 0.93, P=.14, 95% CI -3.19 to 0.44) or quantity (change in mean 15.2 vs 10.6; beta = -4.61, SE = 4.44, P=.30, 95% CI -13.32 to 4.10).

Participants in study arm 1 who did not receive a chat session for whatever reason did reduce their frequency of cannabis use more (change in mean 1.9) than participants in study arm 2 (change in mean 0.7) who did not have the possibility for a chat session due to their allocation (beta = -1.97, SE = 0.60, P=.001, 95% CI -3.14 to -0.80). However, they did not differ with respect

to the reduction in the quantity of cannabis used (change in mean 10.6 vs 10.3; beta = -0.33, SE = 6.48, *P*=.96, 95% CI -13.03 to 12.37). There were no significant differences between these two groups with respect to module completion (mean 2.5, SD 2.1 vs mean 2.9, SD 2.4; *t*=-1.18, df = 159, *P*=.23), but those in study arm 2 showed lower changes in their entries in the consumption diary (change in mean 0.4 vs 0.5; beta = -0.28, SE = 0.12, *P*=.03, 95% CI -0.66 to -0.53) compared to those who did not receive the chat session in study arm 1.

Additional Help and Adverse Events

At the 3-month follow-up, 88.0% of participants (103/117) stated that they had not contacted any other treatment services (7 participants in study arm 1, 2 in study arm 2, and 5 in study arm 3). A total of 5.1% (6/117) had contacted a psychiatrist, 2.6% (3/117) a family doctor, 1.7% (2/117) a psychologist, 1.7% (2/117) a different Internet counseling service, and 1 person (0.9%) a drug counselor. During the whole study period, 5 out of 308 (1.6%) participants contacted one of the outpatient addiction clinics from the ARUD Centers for Addiction Medicine. None of them had to be treated as an emergency case or had to be referred to an inpatient treatment service. Moreover, none of the involved counselors or researchers are aware of any adverse or serious adverse event related to the Can Reduce study that was reported by other addiction counseling services.

Discussion

Principal Findings

The Can Reduce study could reach a different group of cannabis users who do not enter outpatient addiction treatment services. They are older and consume much more cannabis than outpatient service users. The finding that we reached cannabis users with more entrenched problems (eg, daily users) is not consistent with the common perception that those using online interventions have less severe problems than those entering outpatient services. We assume that this finding was most probably due to an age effect. Older users consume longer and possibly also more than younger ones but might feel more stigmatized if they enter an outpatient addiction service, due to their greater responsibilities and roles in social relationships, at work, and in society in general.

Can Reduce participants allocated to the self-help with chat study arm reduced their frequency of cannabis use more than those in the other two arms. Even cannabis abstinence was higher among those who received additional chat counseling relative to those who received self-help only at follow-up. There was a trend (*P*=.06) for a greater reduction in quantity of cannabis use in those who received chat versus those in the waiting list group and only a weak tendency (*P*=.12) for the comparison of those with self-help only versus waiting list. Hence, adding one to two chat counseling sessions that are tailored to the self-help participant data and are based on the same therapy approaches as the self-help part can be worthwhile.

As only one-quarter received at least one chat session, the question arose as to what was actually responsible for the superiority of the self-help with chat study arm. The subgroup analyses showed that those participants in study arm 1 who did not receive a chat session reduced their frequency of cannabis use more than those who received self-help only from the beginning (study arm 2). Thus, even an invitation to a chat session and the knowledge that there is a possibility to have a chat appointment might have improved this main outcome for cannabis use. To the best of our knowledge, there are no similar studies in the literature that have reported a comparable effect. However, our result is in line with the first point of the Supportive Accountability model [39] that argues that human support increases adherence—and potentially outcomes—through accountability to a coach who is seen as trustworthy, benevolent, and having expertise. We took care that our chat counselors were perceived as possessing these attributes in the respective chat study arm.

However, those participants who actually received at least one chat counseling session in study arm 1 still performed better in their reduction of cannabis use and completed more self-help modules than their counterparts who did not receive a chat session in the same study arm. This result is in line with a further point of the Supportive Accountability model [39] expecting better outcomes due to a reciprocal relationship, through which the patient can derive explicit benefits. However, this finding could also be related to a selection bias. Those who actually received at least one chat appointment with their counselor could be a selected group of more compliant and possibly more structured participants who could profit best from their allocated intervention.

If we compare the current results with former studies about the reduction of cannabis use with similar therapeutic approaches, it stands out that participants in the Can Reduce self-help without chat study arm performed worse than those in the Australian *Reduce Your Use* study [15], in which greater effects were achieved in the reduction of the quantity (d=0.06 vs d=0.25 in the Australian sample) and frequency of cannabis use days (d=-0.14 vs d=0.33). This Australian study enrolled cannabis users of a similar age range, but included more females (38.6% vs 24.7%) and users with less severe cannabis consumption at baseline. This may also be the reason that we did not observe greater effects in the Severity of Dependence Scale, in contrast to the Australian study (ITT: d=0.07 vs d=0.33). However, the Australian study provided videos of a real person who provided continuous MI during almost all parts of the intervention. This clearly might have been an advantage compared to our version with only written MI. Another possibility that could potentially increase the engagement of self-help participants might be to provide a personal companion with whom the participants could identify, as we attempted in a similar ongoing trial with problematic cocaine users [40]. The effects in the Can Reduce self-help plus chat study arm were smaller for the quantity of cannabis used (ITT: d=0.09 vs d=0.25) and similar for the frequency of cannabis use days (ITT: d=0.34 vs d=0.33) compared to the Australian self-help trial [15]. Participants in the Can Reduce self-help with chat study arm performed better than those in the more recent German *Quit the Shit* study with respect to the reduction in the frequency of cannabis use days (ITT: d=0.34 vs d=0.20) [11]. The German study recruited younger participants (mean age 24.2 years, SD 5.8 vs mean age 29.8 years, SD 10.0).

We observed a borderline significant effect in the abstention rates between the self-help with chat and the self-help without chat study arms. As we did not initially expect that enough participants would maintain their abstinence, we omitted abstinence as an outcome measure in the study protocol [1]. Abstinence rates were not reported in the German *Quit the Shit* studies [12,14], but comparable differences between study arms with respect to 3 months of abstinence were achieved in this study (8.8%) and in the Australian study (5.8%) [15].

Setting a goal for cannabis consumption was implemented as described in the study protocol [1]. In the introduction to the consumption diary, we recommended that participants should plan to reduce their cannabis use by at least 20 to 30% in the first week and then continue with this strategy in subsequent weeks if they succeeded. For participants who did not succeed, we recommended that they created more modest goals until their final aim was achieved. During the analyses of the consumption diary patterns, we realized that there was a considerable subgroup of participants who preferred to abstain from cannabis even in the first week. Experiences from the chat counseling sessions showed that, although the counselors in the corresponding study arm strengthened this procedure in the self-help intervention part, some participants argued that they had learned from previous experience that they were much more successful in stopping a potentially addictive behavior than in reducing it. In this case, the counselors tried to encourage them to abstain and to assist them in the maintenance of their abstinence. However, this also resulted in some cases with a new challenge. There was a substantial number of participants in this subgroup who very quickly abstained from their cannabis use and who did not log in again, although they were reminded by automated reminder emails and/or their chat counselor, and who then could not be reached at the follow-up assessment. Possible strategies to prevent such early missing cases due to abstention could be specific reminder emails sent automatically and/or by introducing the chat counselor at an earlier stage.

Strengths and Limitations

The strengths of the Can Reduce study are that the intervention is theory based and pretested, that this Web-based intervention was able to reach cannabis users who otherwise would not have sought help, and that we were able to disentangle the effects of chat counseling additional to self-help for the reduction in cannabis use in frequent cannabis users, three-quarters of whom used cannabis daily. This study also possesses limitations that merit consideration. First, we did not biologically validate cannabis consumption for financial reasons, as we did not want to limit participation to participants who were willing to provide, for example, saliva samples, and as we did not want to limit external validity. Second, we did not succeed in attaining a better 6-week follow-up as intended in the study protocol, which limits the explanatory power of the short-term effects of Can Reduce. However, the 3-month follow-up rate (117/308, 38.0%) was comparable to similar studies with problematic cannabis users in Europe [12,14], but rather low compared to Internet-based randomized controlled trials for the improvement of nonaddiction-related problems. Moreover, we used the most reliable imputation method available to handle missing data [38] at follow-up. Third, due to ethical legislations, we had to limit the minimal participation age to 18 years, as younger participants would have needed parental informed consent, and we expected that the overwhelming majority of minors would avoid participation under these conditions. Moreover, this would have been a contradiction with the concept of a maximally anonymous Web-based intervention for the reduction of cannabis use. Cannabis is still illegal in Switzerland and Germany, from where the majority of participants in this study come from. Fourth, participants were randomly allocated into three study arms with slightly different sizes and a block randomization could have prevented this.

Conclusions

In conclusion, the Can Reduce study demonstrated that Web-based interventions possess the potential to reach heavy cannabis users who differ from those who enter outpatient addiction treatment services. We further conclude that offering brief chat counseling in addition to Web-based self-help can significantly increase success in the reduction of cannabis use in the different groups of cannabis users investigated.

Acknowledgments

Funding for this study was provided by Infodrog, the Swiss Office for the Coordination of Addiction Facilities, Switzerland (Grant No. 5012/13/ZH/Cannabis Control). The funding institution had no role in the development or evaluation of the interventions. The authors wish to extend particular appreciation to the psychology master's students Emilija Nikolic and Manja Djordjevic for helping to conduct the telephone follow-up interviews.

Conflicts of Interest

None declared.

Authors' Contributions

MPS was responsible for the study design and the final manuscript. AW and MPS performed the analyses and prepared the first draft of the paper. All authors developed the intervention of study arms 1 and 2 and SH supervised the analyses. AW programmed and implemented the study website, Can Reduce. All of the authors approved the final version of the manuscript submitted for publication.

Multimedia Appendix 1

Can Reduce participant information and informed consent page.

[PDF File (Adobe PDF File), 110KB - jmir_v17i10e232_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.1 [28].

[PDF File (Adobe PDF File), 853KB - jmir_v17i10e232_app2.pdf]

References

1. Schaub MP, Haug S, Wenger A, Berg O, Sullivan R, Beck T, et al. Can reduce--the effects of chat-counseling and web-based self-help, web-based self-help alone and a waiting list control program on cannabis use in problematic cannabis users: a randomized controlled trial. BMC Psychiatry 2013;13:305 [FREE Full text] [doi: 10.1186/1471-244X-13-305] [Medline: 24228630]
2. Thanki D, Matias J, Griffiths P, Noor A, Olszewski D, Simon R, et al. Prevalence of Daily Cannabis Use in the European Union and Norway. Lisbon, Portugal: European Monitoring Centre for Drugs and Drug Addiction; 2013. URL: http://www.emcdda.europa.eu/attachements.cfm/att_191926_EN_emcdda-daily-cannabis-use-2012.pdf [accessed 2015-06-25] [WebCite Cache ID 6ZY4jS0SR]
3. Gmel G, Kuendig H, Notari L, Gmel C. Suchtmonitoring Schweiz - Konsum von Alkohol, Tabak und illegalen Drogen in der Schweiz im Jahr. Lausanne, Switzerland: Sucht Schweiz; 2014. URL: http://www.suchtmonitoring.ch/de/4/1-2.html?cannabis-pravalenz-gebrauch-in-der-gesamtbevolkerung [accessed 2015-06-25] [WebCite Cache ID 6ZY4OgGMy]
4. Fischer B, Jeffries V, Hall W, Room R, Goldner E, Rehm J. Lower Risk Cannabis Use Guidelines for Canada (LRCUG): a narrative review of evidence and recommendations. Can J Public Health 2011;102(5):324-327. [Medline: 22032094]
5. Anthony JC. The epidemiology of cannabis dependence. In: Roffman RA, Stephens RS, editors. Cannabis Dependence: Its Nature, Consequences, Treatment. Cambridge, UK: Cambridge University Press; 2006:58-105.
6. Fischer R, Clair C, Studer J, Cornuz J, Gmel G. Prevalence and factors associated with use of smokeless tobacco in young Swiss men. Eur J Public Health 2014 Jun;24(3):459-464 [FREE Full text] [doi: 10.1093/eurpub/ckt086] [Medline: 23813708]
7. act-info. Jahresbericht 2013: Suchtberatung und Suchtbehandlung in der Schweiz. Ergebnisse des Monitoringsystems. Berne, Switzerland: Federal Office of Public Health; 2015. URL: http://www.suchtschweiz.ch/fileadmin/user_upload/DocUpload/2015_Maffli-Delgrande-Astudillo-etal_BAG-act-info-2013-D.pdf [WebCite Cache ID 6ZY4X0QFh]
8. act-info. Jahresbericht 2011: Suchtberatung und Suchtbehandlung in der Schweiz. Ergebnisse des Monitoringsystems. Berne, Switzerland: Federal Office of Public Health; 2012. URL: http://www.bag.admin.ch/themen/drogen/00042/00632/03290/03295/index.html?lang=de&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCLdIB,g2ym162epYbg2c_JjKbNoKSn6A– [WebCite Cache ID 6ZY4dP9KZ]
9. European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). Treatment of Cannabis-Related Disorders in Europe. Luxembourg City, Luxembourg: Publications Office of the European Union; 2015. URL: http://www.emcdda.europa.eu/attachements.cfm/att_237537_EN_TDXD14017ENN.pdf%20ISSN:2314-9264 [accessed 2015-06-25] [WebCite Cache ID 6ZY4ogrQR]
10. Becker J, Hungerbuehler I, Berg O, Szamrovicz M, Haubensack A, Kormann A, et al. Development of an integrative cessation program for co-smokers of cigarettes and cannabis: demand analysis, program description, and acceptability. Subst Abuse Treat Prev Policy 2013;8:33 [FREE Full text] [doi: 10.1186/1747-597X-8-33] [Medline: 24025478]
11. Curry SJ. eHealth research and healthcare delivery beyond intervention effectiveness. Am J Prev Med 2007 May;32(5 Suppl):S127-S130 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.026] [Medline: 17466817]
12. Tossmann H, Jonas B, Tensil M, Lang P, Strüber E. A controlled trial of an internet-based intervention program for cannabis users. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):673-679. [doi: 10.1089/cyber.2010.0506] [Medline: 21651419]
13. Miller WR, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behavior. New York, NY: Guilford Press; 1991.
14. Jonas B, Tossmann P, Tensil M, Leuschner F, Strüber E. Effekte einer einmaligen chat-intervention auf problematischen substanzkonsum. Sucht 2012 Jun;58(3):173-182 [FREE Full text] [doi: 10.1024/0939-5911.a000182]
15. Rooke S, Copeland J, Norberg M, Hine D, McCambridge J. Effectiveness of a self-guided Web-based cannabis treatment program: randomized controlled trial. J Med Internet Res 2013;15(2):e26 [FREE Full text] [doi: 10.2196/jmir.2256] [Medline: 23470329]
16. Goldstein MG, Niaura R, Follick MJ, Abrams DB. Effects of behavioral skills training and schedule of nicotine gum administration on smoking cessation. Am J Psychiatry 1989 Jan;146(1):56-60. [Medline: 2643357]
17. Steinberg KL, Roffman R, Carroll K, McRee B, Babor T, Miller M, et al. Brief Counseling for Marijuana Dependence: A Manual for Treating Adults. Rockville, MD: Center for Substance Abuse Treatment, Substance Abuse and Mental Health

Services Administration, US Department of Health and Human Services; 2005. URL: http://store.samhsa.gov/shin/content//SMA12-4211/SMA12-4211.pdf [accessed 2015-10-06] [WebCite Cache ID 6c5QLM5I5]
18. Copeland J, Swift W, Roffman R, Stephens R. A randomized controlled trial of brief cognitive-behavioral interventions for cannabis use disorder. J Subst Abuse Treat 2001 Sep;21(2):55-64; discussion 65. [Medline: 11551733]
19. Blankers M, Koeter MW, Schippers GM. Internet therapy versus internet self-help versus no treatment for problematic alcohol use: a randomized controlled trial. J Consult Clin Psychol 2011 Jun;79(3):330-341 [FREE Full text] [doi: 10.1037/a0023498] [Medline: 21534652]
20. Can Reduce. URL: http://www.canreduce.ch/ [accessed 2015-09-29] [WebCite Cache ID 6bulmu9lx]
21. Schaub M, Sullivan R, Stark L. Snow control - an RCT protocol for a Web-based self-help therapy to reduce cocaine consumption in problematic cocaine users. BMC Psychiatry 2011;11:153 [FREE Full text] [doi: 10.1186/1471-244X-11-153] [Medline: 21943294]
22. Schaub M, Sullivan R, Haug S, Stark L. Web-based cognitive behavioral self-help intervention to reduce cocaine consumption in problematic cocaine users: randomized controlled trial. J Med Internet Res 2012;14(6):e166 [FREE Full text] [doi: 10.2196/jmir.2244] [Medline: 23192752]
23. Drugs Meter. URL: http://www.drugsmeter.com/ [accessed 2015-09-30] [WebCite Cache ID 6bw7nTJGf]
24. Becker J, Haug S, Kraemer T, Schaub MP. Feasibility of a group cessation program for co-smokers of cannabis and tobacco. Drug Alcohol Rev 2015 2015;34(4):418-426 [FREE Full text] [Medline: 25676414]
25. Karschner EL, Darwin WD, McMahon RP, Liu F, Wright S, Goodwin RS, et al. Subjective and physiological effects after controlled Sativex and oral THC administration. Clin Pharmacol Ther 2011 Mar;89(3):400-407. [doi: 10.1038/clpt.2010.318] [Medline: 21289620]
26. Schaub M, Stohler R. Swiss Medical Forum. 2006. Long-term adverse effects of cannabis - an underestimated problem? URL: http://www.medicalforum.ch/docs/smf/archiv/de/2006/2006-50/2006-50-060.pdf [accessed 2015-06-25] [WebCite Cache ID 6ZY50OkTJ]
27. Schaub MP, Uchtenhagen A. Building a European consensus on minimum quality standards for drug treatment, rehabilitation and harm reduction. Eur Addict Res 2013;19(6):314-324. [doi: 10.1159/000350740] [Medline: 23774998]
28. Schaub M, Wenger A, Bachmann A, Stark L, Haug S. Quality assurance in the guided Internet-based addiction aid as exemplified by the Swiss addiction portal Safe Zone. Suchtmedizin Forsch und Prax 2004;16(4):187-193.
29. Eysenbach G. CONSORT-EHEALTH: implementation of a checklist for authors and editors to improve reporting of Web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]
30. Annaheim B, Scotto TJ, Gmel G. Revising the Cannabis Use Disorders Identification Test (CUDIT) by means of Item Response Theory. Int J Methods Psychiatr Res 2010 Sep;19(3):142-155. [doi: 10.1002/mpr.308] [Medline: 20812291]
31. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
32. Gossop M, Darke S, Griffiths P, Hando J, Powis B, Hall W, et al. The Severity of Dependence Scale (SDS): psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. Addiction 1995 May;90(5):607-614. [Medline: 7795497]
33. Allsop DJ, Norberg MM, Copeland J, Fu S, Budney AJ. The Cannabis Withdrawal Scale development: patterns and predictors of cannabis withdrawal and distress. Drug Alcohol Depend 2011 Dec 1;119(1-2):123-129. [doi: 10.1016/j.drugalcdep.2011.06.003] [Medline: 21724338]
34. Schnell T, Münchenhagen L, Tersudi K, Daumann J, Gouzoulis-Mayfrank E. Entwicklung und evaluation eines deutschsprachigen instruments zur erfassung von cannabis-craving (CCS-7). Zeitschrift für Klinische Psychologie und Psychotherapie 2011 Jan;40(1):33-41 [FREE Full text] [doi: 10.1026/1616-3443/a000072]
35. Heishman SJ, Singleton EG, Liguori A. Marijuana Craving Questionnaire: development and initial validation of a self-report instrument. Addiction 2001 Jul;96(7):1023-1034. [doi: 10.1080/09652140120053084] [Medline: 11440613]
36. Kokkevi A, Hartgers C. EuropASI: European adaptation of a multidimensional assessment instrument for drug and alcohol dependence. Eur Addict Res 1995;1(4):208-210 [FREE Full text] [doi: 10.1159/000259089]
37. Rumpf HJ, Meyer C, Hapke U, John U. Screening for mental health: validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard. Psychiatry Res 2001 Dec 31;105(3):243-253. [Medline: 11814543]
38. 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]
39. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
40. Schaub MP, Maier LJ, Wenger A, Stark L, Berg O, Beck T, et al. Evaluating the efficacy of a Web-based self-help intervention with and without chat counseling in reducing the cocaine use of problematic cocaine users: the study protocol of a pragmatic three-arm randomized controlled trial. BMC Psychiatry 2015;15:156 [FREE Full text] [doi: 10.1186/s12888-015-0518-6] [Medline: 26159624]

Abbreviations

act-info: addiction, care and therapy information
ANOVA: analysis of variance
ARUD: Arud Centers for Addiction Medicine
BSM: behavioral self-management
CBT: cognitive behavioral therapy
CCS-7: Cannabis Craving Symptoms questionnaire
CUDIT: Cannabis Use Disorders Identification Test
CWS: Cannabis Withdrawal Scale
FDA: Fragebogen Substanzanamnese (questionnaire for the assessment of substance use history)
ISGF: Swiss Research Institute for Public Health and Addiction
ITT: intention-to-treat
MHI-5: Mental Health Inventory
MI: motivational interviewing
N/A: not applicable
OR: odds ratio
RCT: randomized controlled trial
SDS: Severity of Dependence Scale
WL: waiting list

Edited by G Eysenbach; submitted 06.07.15; peer-reviewed by R Tait, B Jonas, U Buchner, N Arnaud, D Reinwand; comments to author 29.07.15; revised version received 27.08.15; accepted 21.09.15; published 13.10.15.

Please cite as:
Schaub MP, Wenger A, Berg O, Beck T, Stark L, Buehler E, Haug S
A Web-Based Self-Help Intervention With and Without Chat Counseling to Reduce Cannabis Use in Problematic Cannabis Users: Three-Arm Randomized Controlled Trial
J Med Internet Res 2015;17(10):e232
URL: http://www.jmir.org/2015/10/e232/
doi:10.2196/jmir.4860
PMID:26462848

Original Paper

Effects of a Web-Based Personalized Intervention on Physical Activity in European Adults: A Randomized Controlled Trial

Cyril FM Marsaux[1], MSc; Carlos Celis-Morales[2], PhD; Rosalind Fallaize[3], PhD; Anna L Macready[3], PhD; Silvia Kolossa[4], MSc; Clara Woolhead[5], MSc; Clare B O'Donovan[5], MSc; Hannah Forster[5], BSc; Santiago Navas-Carretero[6,7], PhD; Rodrigo San-Cristobal[6], MSc; Christina-Paulina Lambrinou[8], MSc; George Moschonis[8], PhD; Agnieszka Surwillo[9], MSc; Magdalena Godlewska[9], MSc; Annelies Goris[10], PhD; Jettie Hoonhout[11], PhD; Christian A Drevon[12], MD, PhD; Yannis Manios[8], PhD; Iwona Traczyk[9], PhD; Marianne C Walsh[5], PhD; Eileen R Gibney[5], PhD; Lorraine Brennan[5], PhD; J Alfredo Martinez[6,7], PhD; Julie A Lovegrove[3], PhD; Michael J Gibney[5], PhD; Hannelore Daniel[4], PhD; John C Mathers[2], PhD; Wim HM Saris[1], MD, PhD

[1]Department of Human Biology, School of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University Medical Centre + (MUMC+), Maastricht, Netherlands

[2]Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, United Kingdom

[3]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, University of Reading, Reading, United Kingdom

[4]Zentralinstitut für Ernährungs- und Lebensmittelforschung (ZIEL) Research Center of Nutrition and Food Sciences, Biochemistry Unit, Technische Universität München, München, Germany

[5]University College Dublin (UCD) Institute of Food and Health, University College Dublin, Belfield, Dublin, Ireland

[6]Department of Nutrition, Food Science and Physiology, Centre for Nutrition Research, University of Navarra, Pamplona, Spain

[7]Centro de Investigación Biomédica en Red-Fisiopatología de la Obesidad y Nutrición (CIBERobn), Instituto de Salud Carlos III, Madrid, Spain

[8]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece

[9]National Food & Nutrition Institute (IZZ), Warsaw, Poland

[10]Personal Health Solutions, Philips Consumer Lifestyle, Amsterdam, Netherlands

[11]Experiences Research Department, Philips Research, Eindhoven, Netherlands

[12]Department of Nutrition, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Cyril FM Marsaux, MSc
Department of Human Biology
School of Nutrition and Translational Research in Metabolism (NUTRIM)
Maastricht University Medical Centre + (MUMC+)
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433881379
Fax: 31 433670976
Email: c.marsaux@maastrichtuniversity.nl

Abstract

Background: The high prevalence of physical inactivity worldwide calls for innovative and more effective ways to promote physical activity (PA). There are limited objective data on the effectiveness of Web-based personalized feedback on increasing PA in adults.

Objective: It is hypothesized that providing personalized advice based on PA measured objectively alongside diet, phenotype, or genotype information would lead to larger and more sustained changes in PA, compared with nonpersonalized advice.

Methods: A total of 1607 adults in seven European countries were randomized to either a control group (nonpersonalized advice, Level 0, L0) or to one of three personalized groups receiving personalized advice via the Internet based on current PA plus diet (Level 1, L1), PA plus diet and phenotype (Level 2, L2), or PA plus diet, phenotype, and genotype (Level 3, L3). PA was measured for 6 months using triaxial accelerometers, and self-reported using the Baecke questionnaire. Outcomes were objective and self-reported PA after 3 and 6 months.

Results: While 1270 participants (85.81% of 1480 actual starters) completed the 6-month trial, 1233 (83.31%) self-reported PA at both baseline and month 6, but only 730 (49.32%) had sufficient objective PA data at both time points. For the total cohort after 6 months, a greater improvement in self-reported total PA (P=.02) and PA during leisure (nonsport) (P=.03) was observed in personalized groups compared with the control group. For individuals advised to increase PA, we also observed greater improvements in those two self-reported indices (P=.006 and P=.008, respectively) with increased personalization of the advice (L2 and L3 vs L1). However, there were no significant differences in accelerometer results between personalized and control groups, and no significant effect of adding phenotypic or genotypic information to the tailored feedback at month 3 or 6. After 6 months, there were small but significant improvements in the objectively measured physical activity level (P<.05), moderate PA (P<.01), and sedentary time (P<.001) for individuals advised to increase PA, but these changes were similar across all groups.

Conclusions: Different levels of personalization produced similar small changes in objective PA. We found no evidence that personalized advice is more effective than conventional "one size fits all" guidelines to promote changes in PA in our Web-based intervention when PA was measured objectively. Based on self-reports, PA increased to a greater extent with more personalized advice. Thus, it is crucial to measure PA objectively in any PA intervention study.

Trial Registration: ClinicalTrials.gov NCT01530139; http://clinicaltrials.gov/show/NCT01530139 (Archived by WebCite at: http://www.webcitation.org/6XII1QwHz)

(J Med Internet Res 2015;17(10):e231) doi:10.2196/jmir.4660

KEYWORDS

physical activity; eHealth; randomized controlled trial; personalized nutrition; genotype; phenotype; Internet

Introduction

Physical inactivity is one of the major risk factors for noncommunicable diseases [1]. It has been estimated that in 2008, approximately 7.3% of the 9.2 million deaths occurring in Europe were attributed to physical inactivity compared with 3.7% attributed to obesity [2]. Increasing physical activity (PA) continues to be a public health priority. Although public knowledge of the health benefits of regular PA is good and the recommendation of "30 min per day of activity most days of the week" is recognized widely, recent data from the World Health Organization (WHO) suggest that 35% of European adults do not meet PA recommendations [3].

Finding effective ways to increase PA is challenging. The limited success of "one size fits all" PA promotion programs may be partly due to the fact that inactive individuals are unaware that their current PA is inadequate [4,5]. Thus, providing personalized PA feedback may be more effective in increasing PA than a nonpersonalized conventional approach. Internet-based interventions for PA may have potential to increase levels of PA because large numbers of physically inactive individuals can be reached. However, it has been pointed out that positive effects were quite small, and might not be sustained in the long term. Furthermore, objective PA measurements and greater sample sizes are required [6-9].

Although many studies have used self-reports, such as the Baecke questionnaire or the International Physical Activity Questionnaire (IPAQ) [10], objective measurements of PA are more reliable [11]. Developments of PA measurement devices in the last decade have improved quantification of PA in free-living subjects. Accelerometers, for instance, have become popular because they can be worn without major inconvenience, require little effort by the user, and are compatible with most daily activities. Because they can record data for up to several weeks and measure PA accurately [12,13], they are useful research tools. In spite of this, few studies comparing tailored PA advice with nontailored or no advice have included these objective measures of PA [14-17].

Whether personalized PA feedback promotes behavioral change remains unclear. We used data collected during the Food4Me Study, which was registered at ClinicalTrials.gov (NCT01530139), to investigate the impact of different levels of personalization on PA change, using phenotypic and genotypic information to tailor the PA advice (see also Multimedia Appendix 1). We hypothesized that individually tailored advice would lead to greater and more sustained changes in PA, and that the intervention would be more effective as the level of personalization increased.

Methods

Study Design

Full details of the study protocols have been described elsewhere [18]. Briefly, the Food4Me proof-of-principle study was a 6-month, 4-arm, randomized controlled trial (RCT) conducted across seven European countries to compare the effects of three levels of personalized advice with standard population advice on health-related outcomes. The intervention was designed to emulate an Internet-based service [19] and aimed to answer the following primary questions: (1) does personalization of dietary and PA advice result in bigger improvement in diet and PA compared with nonpersonalized, conventional guidelines? And (2) is personalization based on individualized phenotypic or genotypic information more effective in assisting and/or motivating participants to make and sustain appropriate healthy changes than personalization based on analysis of baseline diet and PA alone? To answer these questions, participants were randomized to a control group (Level 0) or to one of three personalized intervention groups with increasingly more detailed personalized advice (Levels 1 to 3) for 6 months. The levels are described in Textbox 1.

Textbox 1. Description of control and intervention groups and their levels of personalization in the Food4Me Study.

Levels of personalization

- Level 0 (L0; control group): nonpersonalized advice based on (European) general guidelines for diet and PA
- Level 1 (L1): personalized advice based on individual dietary intake and PA alone
- Level 2 (L2): personalized advice based on individual dietary intake, PA, and phenotypic data
- Level 3 (L3): personalized advice based on individual dietary intake, PA, and phenotypic and genotypic data

In the personalized groups, personalization of the PA advice was greater in L2 and L3 compared with L1: it was linked to phenotypic data (waist circumference, blood total cholesterol; L2 and L3) and genotypic data (fat mass- and obesity-associated gene, *FTO*; L3). See also the section Physical Activity Feedback below.

Outcomes

We focus here on PA after 3 and 6 months of intervention. PA is presented using both objective (accelerometer) and self-reported (PA questionnaire) data. Other outcomes included dietary intake, but are not within the scope of this paper.

Recruitment

We aimed to recruit 1540 participants aged ≥18 years in seven European countries—Germany, Greece, Ireland, the Netherlands, Poland, Spain, and the United Kingdom [18]. Subjects were ineligible to take part in the study if they had no or limited access to the Internet, were following a prescribed diet, or had altered nutritional requirements because of a medical condition. The ethics committee from each recruiting center approved the study protocol. Between August 2012 and August 2013, 1607 adults (653 men, 40.63%; and 954 women, 59.37%) were randomized to the intervention. All participants gave informed consent digitally before participating in the study.

Measures

Data were collected using standard operating procedures in all seven countries [18]. Participants received study kits by post, containing all necessary materials (including an accelerometer) to perform measurements at home, but used their own scales to measure body weight. Printed and digital instructions, as well as online videos, were available for all participants in the languages of all seven countries.

Objectively Measured Physical Activity

Objective Physical Activity Monitoring

Habitual PA was assessed objectively using the TracmorD triaxial accelerometer (Philips Consumer Lifestyle, the Netherlands) [20,21]. The device is small (3.2 × 3.2 × 0.5 cm), light (12.5 g), waterproof to a depth of 30 m, has a battery life of 3 weeks, and can record data for up to 22 weeks.

Participants activated the TracmorD accelerometer by creating an account online, installing an app on their computer, and connecting the device to the computer using the USB adapter provided. Upon activation, men could choose between three wearing positions—pocket, belt, or necklace—and women between four wearing positions—pocket, belt, necklace, or bra. Participants could change their wearing position after informing the research team, who would update the position in the online system. Participants were instructed to wear the accelerometer every day during waking hours, except when taking a shower, for the entire duration of the study. Participants uploaded data every 2 weeks by connecting their monitor to their computer. Researchers checked this regularly and sent reminders to participants, if necessary. The data were transferred in real time and stored on a secured server.

Objective Physical Activity Data Processing

Data were recorded with a time-sampling interval of 1 minute. A day was considered valid if the participant had worn the TracmorD for at least 10 hours but not longer than 18 hours. Wear time was defined as 24 hours minus nonwear time. To define nonwear time, we adapted the recommendations of Choi et al [22] to the TracmorD. Physical activity level (PAL)—the ratio of total energy expenditure to basal metabolic rate—data per minute were estimated from activity counts [20]. Nonwear time was defined by an interval of at least 90 consecutive minutes of PAL per minute values below 1.3889, allowing for 2-minute intervals of values above the threshold, with the upstream or downstream 30 consecutive values below the threshold (for detection of artifactual movements). R software version 3.1.2 [23] was used for all data handling.

Objective Physical Activity Variables

Daily PAL, activity energy expenditure (AEE), and time spent in different PA intensities were derived from the accelerometer.

PAL-per-day calculations were based on the work by Bonomi et al [20]:

$$PAL = 1.354 + (256 \times 10^{-9}) \times counts_{day} \quad (1)$$

where $counts_{day}$ are the sum of minute-by-minute activity counts over 24 hours. However, the wearing position is taken into account, using the belt position as a reference and applying a correction factor for the other positions.

AEE per day was calculated as follows:

$$AEE = (0.9 \times PAL\ per\ day - 1) \times BMR \quad (2)$$

where the daily basal metabolic rate (BMR) is estimated using the Oxford equations developed by Henry, based on sex, age, and weight [24].

Classification into sedentary, and light-, moderate-, and vigorous-intensity PA (LPA, MPA, and VPA, respectively) was based on the application of thresholds for AEE: 0.025, 0.05, and 0.1 kcal/(kg×min) corresponding to 1.5, 3, and 6 metabolic equivalents (METs), respectively. Sedentary time and time spent in LPA, MPA, and VPA were determined by summing the time

during which AEE per minute met the criterion for the appropriate intensity. Finally, moderate-equivalent PA was defined as follows:

Moderate-equivalent PA = MPA + 2 × VPA (3)

to account for the fact that 1 minute of VPA is equivalent to 2 minutes of MPA [25]. Moderate-equivalent PA duration data were also calculated for activity occurring in modified bouts of 10 minutes (ie, with allowance for interruptions of ≤2 minutes at a lower PA intensity) [26].

PA estimates at baseline, month 3, and month 6 were calculated over a 2-week period at each time point. This 2-week assessment period occurred before any feedback was given to participants. Sufficient PA data at each time point was defined as having at least 3 valid weekdays and 2 valid weekend days of accelerometer wear during the 2-week assessment period. For individuals with sufficient PA data, mean data per day were calculated for all objective PA variables using all valid week and weekend days of the assessment period as follows:

Mean = (mean for weekdays × 5 + mean for weekend days × 2) / 7 (4).

For sedentary time and time spent in LPA, MPA, VPA, and moderate-equivalent PA, weekly estimates were also calculated as follows:

Mean = (mean for weekdays × 5 + mean for weekend days × 2) (5).

Self-Reported Physical Activity

At baseline, month 3, and month 6, participants completed the Baecke questionnaire online [27] based on their PA during the last month. The Baecke questionnaire is a short, validated questionnaire assessing habitual PA according to the context in which it occurs and is organized into three sections: (1) PA at work, (2) sport, and (3) during leisure time excluding sport [27-29]. Indices for these three PA categories—work index, sport index, and leisure time (nonsport) index, each ranging from 1 to 5—as well as a total activity index—sum of the three previous indices, ranging from 3 to 15—were calculated according to the questionnaire protocol [27].

Anthropometrics

Participants self-measured their height, weight, and waist circumference, and uploaded their measurements directly onto their personal Food4Me online account [18]. Validation of self-reported sociodemographic and anthropometric measures have been described elsewhere [30].

Genotyping

Participants collected a buccal cell sample at baseline, using Isohelix SK-1 DNA buccal swabs and Isohelix Dri-capsules (LGC Genomics, Hertfordshire, UK). Samples were returned to their recruiting center and shipped to LGC Genomics, who extracted the DNA and used competitive allele-specific polymerase chain reaction (KASP) genotyping assays to provide biallelic scoring of single nucleotide polymorphism (SNP) rs9939609 in the *FTO* gene.

Physical Activity Feedback

All participants received a feedback report at months 0 and 3 via email in PDF. Reports were also available on the participant's personal Food4Me account. Participants in personalized groups (L1, L2, and L3) received personalized feedback based on accelerometer data (or self-reported data if accelerometer data were not available), whereas controls (L0) received a PDF containing general guidelines (see the following section). At each time point, researchers calculated the average PAL based on 2 weeks of accelerometer data collection for each participant and used it in the derivation of the PA feedback for personalized groups (L1 to L3). The feedback report was sent 1 to 2 weeks after this 2-week measuring period. For L0, the same generalized advice was sent at months 0 and 3. After completing the study, all participants received a personalized report based on the 6-month intervention.

Cutoffs Definition

PA was defined as adequate if objective PAL was ≥1.8. A value of PAL ≥1.5 and <1.8 was considered low and a PAL of <1.5 was considered very low. If accelerometer data were not available at the time of feedback, the total activity index of the Baecke questionnaire was used instead of PAL with the following cutoffs: ≥8.5 (adequate PA), ≥5.5 to <8.5 (low), and <5.5 (very low).

Derivation of Feedback Messages in Relation to Physical Activity

Level 0 (Controls)

Participants in the control group received the nonpersonalized advice that they should be physically active at least 150 minutes per week.

Feedback reports in personalized groups contained specific messages, selected according to standardized algorithms, based on subjects' characteristics and their allocated group.

Level 1

In the PA section of the personalized report, current level of PA was indicated with a mark on a three-color line, based on the cutoffs defined above: red area (very low PAL), amber (low PAL), and green (adequate PAL). The report included tailored advice to increase strongly, increase, or maintain PA based on current PAL and body mass index (BMI), as well as tips on how to be (more) physically active. Participants had access to additional information about PA and tips online on their personal Food4Me account. Hyperlinks to this section of the website were included in the tailored report and participants were encouraged to visit the webpage.

Level 2

Participants in L2 had access to the same information as those in L1. However, the specific PA message in the personalized report was based on current PAL and BMI as well as on individuals' waist circumference and blood total cholesterol.

Level 3

Participants in L3 had access to the same information as those in L1 and L2, in addition to whether they carried the risk allele

(A) for the *FTO* gene; this information was included in the specific PA message alongside current PAL, BMI, waist circumference, and blood total cholesterol. For example, an inactive obese L3 participant with *FTO* risk (AA or AT genotype), high waist circumference, and high total cholesterol would receive the following advice:

> *Your BMI is greater than the recommended healthy range (...). Your waist circumference is also higher than recommended (...). We recommend reducing your body weight and waist circumference to a healthy normal range because you have a genetic variation that can benefit by reducing these two obesity markers (...). Also, your physical activity level is too low; improving your physical activity level will help you to reduce your weight. Your fasting cholesterol level was above the recommended level and we advise you to go to the G.P. [general practitioner] to get this re-checked (...). Become more physically active; to maintain weight loss, 60-90 minutes of moderately intense aerobic activities, such as brisk walking, swimming or cycling, on most days of the week, is recommended. This will also help to lower cholesterol levels.*

PA feedback for L3 participants not carrying the risk for *FTO* was similar to that of L2 participants. However, L3 participants all received information on both *FTO* and 4 other diet-related genetic variants, whereas L2 participants did not receive any genetic-based information. More details of the feedback reports and the Food4Me website are given in the supplementary material (see Multimedia Appendix 2) and elsewhere [18].

Statistical Analysis

Data were analyzed on an intention-to-treat basis. We defined 3 orthogonal contrasts to answer our research questions: first comparing L0 with L1 to L3, then L1 with L2 and L3, and finally L2 with L3. More specifically, to answer the first research question—Is personalized advice more effective than the conventional one size fits all?—intervention effects on PA variables were assessed. We used robust multiple linear regression analysis, based on computation of SMDM estimates [31] to account for violation of the normality assumption, with baseline PA variable, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight as covariates. For accelerometer-derived PA variables, change in accelerometer wear time was included as an additional covariate. The principal assessment of intervention used Contrast 1 comparing L0 with the mean of L1 to L3. First, a generic approach was used where intervention effects for the total cohort were investigated. Second, a targeted approach was used in which the intervention effects on PA, only for participants who received advice to increase their PA, were investigated. For the second part, outcomes for those who received tailored advice targeting PA were compared with the subset of matched L0 (control) participants (ie, controls who would have received personalized advice to increase PA if they had been in a personalized group instead of L0). These matched L0 participants were selected by applying the same algorithm used for individuals in personalized groups.

The second research question—"Is personalization based on individualized dietary, phenotypic, or genotypic information more effective in promoting changes in PA than personalization based on diet and PA alone?"—was tested using two other contrasts. For Contrast 2, comparison of L1 with L2 and L3 tested whether personalization based on phenotypic and/or genotypic information differed from that based on dietary and PA assessment only. For Contrast 3, comparison of L2 with L3 tested whether the addition of genotypic information promoted a greater increase in PA than when using phenotypic, dietary, and PA information only. The analyses outcomes were the same PA variables as for Contrast 1 and both generic and targeted approaches were used.

Sensitivity analyses were performed to compare dropouts with completers and noncompliant (ie, those with too few valid days) with compliant participants. These analyses were performed using robust multiple linear regression for continuous variables and logistic regression for categorical variables, adjusting for sex, age, and country as well as baseline accelerometer wear time and season for accelerometer variables. When examining differences in BMI between dropouts and completers, screening data were used rather than baseline data because 38% of dropouts had no baseline data. R software version 3.1.2 [23] was used to perform all analyses and the significance level was set at $P<.05$.

Results

Study Participants

Of the 5562 individuals who expressed an interest in the Food4Me Study between August 2012 and August 2013, 4044 (72.71%) completed the whole screening process (Figure 1). Of those, 2764 (68.35%) were eligible to take part in the intervention study. The first 1607 of the 2764 (58.14%) participants were randomized to one of the four intervention arms and 127 (7.90%) dropped out immediately after randomization (Figure 1).

The characteristics of the 1480 participants who started the trial and completed baseline measurements are given by intervention arm in Table 1 and Multimedia Appendix 3, and are described elsewhere [18]. Overall, 58.45% (865/1480) were women, the mean age was 39.9 (SD 13.0) years, and 46.15% (683/1480) of participants were overweight or obese. Mean PAL was 1.73 (SD 0.18), and participants spent on average 12.4 (SD 1.3) h/d in sedentary behaviors and 57 (SD 45) min/d in moderate-equivalent PA (29 [SD 32] min/d in modified 10-minute bouts). Mean self-reported total activity index was 7.80 (SD 1.48) (Table 1). Of the 371 participants in L3, 257 (69.3%) and 113 (30.5%) were carriers of the risk (AA or AT) and nonrisk (TT) genotypes for *FTO*, respectively (Table 1), and were therefore informed that they had or did not have the risk variant. A total of 807 of 1120 (72.05%) individuals randomized to the personalized groups (L1 to L3) were not sufficiently active based on baseline measurements and therefore received feedback that they should increase their level of PA (data not shown).

Table 1. Baseline characteristics[a] of the Food4Me participants.

Variables	Control	Personalized advice		
	Level 0 (L0) (n=360)	Level 1 (L1) (n=373)	Level 2 (L2) (n=376)	Level 3 (L3) (n=371)
Ethnicity (white), n (%)	344 (95.6)	363 (97.3)	368 (97.9)	357 (96.2)
Sex (women), n (%)	213 (59.2)	212 (56.8)	220 (58.5)	220 (59.3)
Age (years), mean (SD)	39.5 (13.3)	39.7 (12.9)	40.2 (12.8)	40.2 (13.1)
Anthropometrics				
Height (cm), mean (SD)	171.3 (9.4)	171.3 (9.5)	170.7 (9.4)	171.2 (9.5)
Weight (kg), mean (SD)	74.6 (15.5)	74.1 (16.6)	74.9 (15.9)	75.5 (15.5)
Body mass index (kg/m^2), mean (SD)	25.4 (4.7)	25.2 (5.0)	25.6 (4.9)	25.7 (4.8)
Overweight, n (%)	119 (33.1)	96 (25.7)	103 (27.4)	131 (35.3)
Obese, n (%)	52 (14.4)	57 (15.3)	70 (18.6)	55 (14.8)
Current smokers, n (%)	49 (13.6)	44 (11.8)	34 (9.0)	47 (12.7)
Ex-smokers, n (%)	88 (24.4)	98 (26.3)	101 (26.9)	91 (24.5)
Nonsmokers, n (%)	223 (61.9)	231 (61.9)	241 (64.1)	233 (62.8)
***FTO*[b] genotype, n (%)**				
AA	60 (16.7)	69 (18.5)	66 (17.6)	69 (18.6)
AT	187 (51.9)	175 (46.9)	189 (50.3)	188 (50.7)
TT	112 (31.1)	127 (34.0)	117 (31.1)	113 (30.5)
Objective physical activity (PA) (n_{L0}=303, n_{L1}=324, n_{L2}=339, n_{L3}=321), mean (SD)				
Physical activity level	1.71 (0.18)	1.75 (0.21)	1.73 (0.16)	1.74 (0.17)
Activity energy expenditure (kcal/d)	832 (269)	896 (312)	869 (274)	874 (283)
Sedentary time (min/d)	746 (76)	738 (75)	747 (78)	749 (77)
Light-intensity PA (min/d)	70 (27)	76 (33)	74 (31)	76 (30)
Moderate-intensity PA (min/d)	30 (19)	35 (20)	33 (21)	34 (22)
Vigorous-intensity PA (min/d)	11 (16)	14 (20)	11 (14)	10 (14)
Moderate-equivalent PA[c] (min/d)	53 (43)	63 (50)	56 (42)	55 (45)
Moderate-equivalent PA in bouts (min/d)	27 (32)	34 (38)	28 (28)	28 (30)
Self-reported PA (n_{L0}=359, n_{L1}=371, n_{L2}=376, n_{L3}=370), mean (SD)				
Total activity index	7.71 (1.47)	7.94 (1.48)	7.78 (1.43)	7.80 (1.54)
Work index	2.26 (0.59)	2.30 (0.61)	2.28 (0.62)	2.29 (0.62)
Sport index	2.70 (0.87)	2.85 (0.89)	2.73 (0.85)	2.74 (0.88)
Leisure time (nonsport) index	2.75 (0.69)	2.81 (0.70)	2.78 (0.69)	2.78 (0.67)

[a]Data are presented as unadjusted means (SD) for continuous variables and absolute numbers (%) for categorical variables. Levels 1 to 3 received personalized advice; only participants in Level 3 were informed whether they carried or did not carry the risk allele for *FTO* (A).

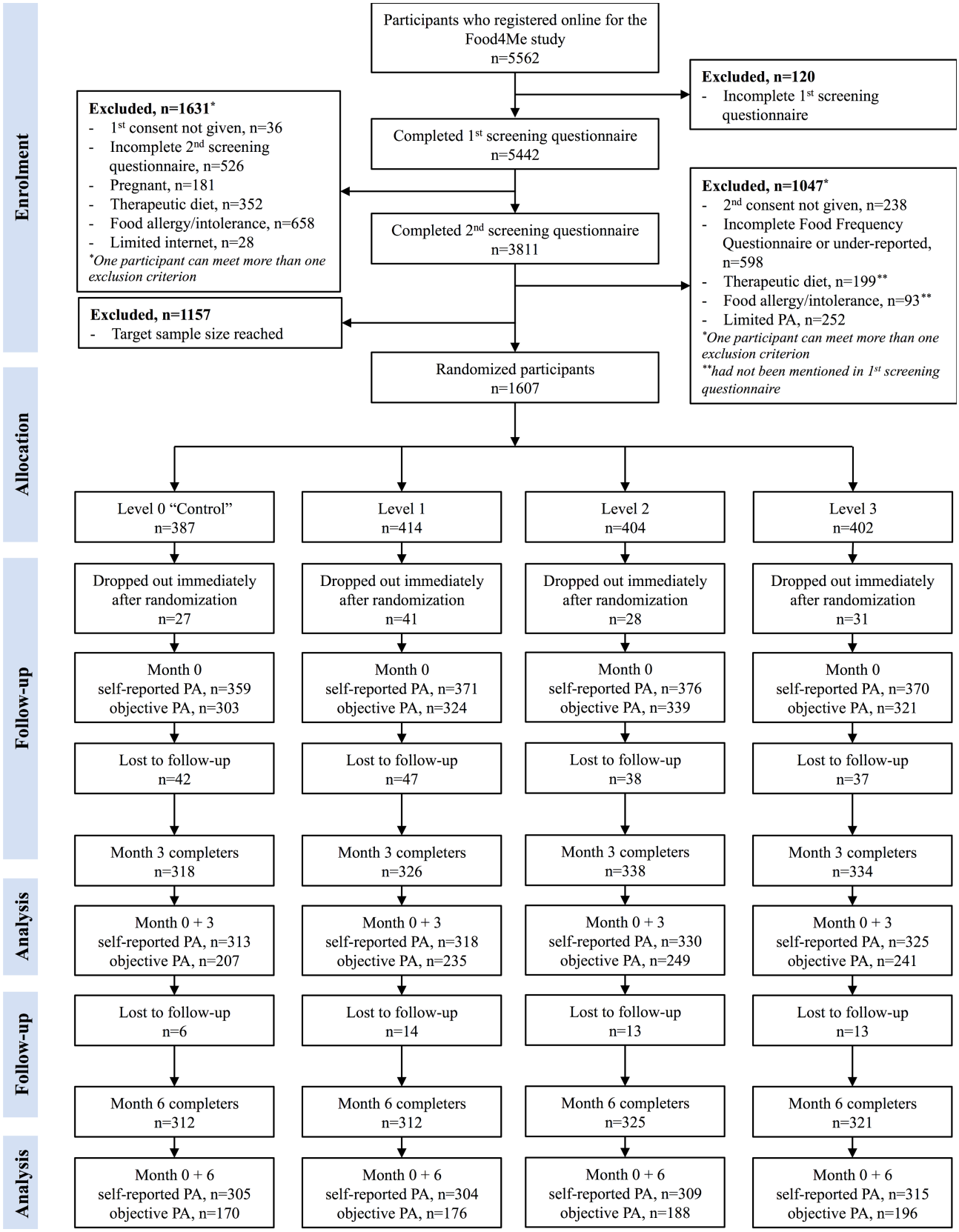
[b]Fat mass and obesity associated (*FTO*).

[c]Moderate-equivalent PA is (MPA + 2 × VPA).

Of the 1607 randomized participants, 1270 (79.03%) completed the 6-month intervention and 1233 (76.73%) had self-reported data on PA for both baseline and month 6, whereas only 730 (45.43%) had sufficient valid accelerometer data for both time points (Figure 1). Dropouts were more likely to be women (odds ratio [OR] 1.34, 95% CI 1.05-1.75, *P*=.03), were younger than completers (*P*<.001), and had a higher BMI at screening than completers (*P*=.02). Smoking habits did not differ significantly between dropouts and completers, and the dropout rate was similar in all four groups, L0 to L3 (data not shown). Among completers, those who were not compliant with objective PA measurement at month 6 were younger (*P*<.001), had a higher

baseline BMI (P=.04), and a lower baseline PAL (P=.03). There were no significant differences in smoking habits between those who were compliant and those who were not, and compliance was similar for men and women and in all four groups, L0 to L3 (data not shown).

Figure 1. Flow of participants through the study. PA: physical activity.



Effect of Different Levels of Personalized Advice on Objective Physical Activity

Total Cohort: Generic Approach

At month 3 (see Multimedia Appendix 4), participants increased their PAL (L0: +0.02, *P*=.008 and L1 to L3: +0.02, *P*<.001) and AEE (L0: +24.2 kcal/d, *P*=.03 and L1 to L3: +19.5 kcal/d, *P*=.001), and spent significantly more time in MPA (L0: +18 min/wk, *P*=.01 and L1 to L3: +17 min/wk, *P*<.001) and less time in sedentary behavior (L0: -148 min/wk, *P*<.001 and L1 to L3: -133 min/wk, *P*<.001). No significant change in PA was observed at month 6 (see Multimedia Appendix 5), except for a significant decrease in sedentary time (L0: -190 min/wk, *P*<.001 and L1 to L3: -155 min/wk, *P*<.001). Furthermore, we found no significant differences in objectively measured PA between control and personalized groups, or between personalized groups, at month 3 or 6 (Multimedia Appendices 4 and 5).

Participants Who Received Advice to Increase Physical Activity and Matched Controls: Targeted Approach

At month 3, we observed significant improvements in all components of PA for participants in personalized groups as well as for matched controls (see Multimedia Appendix 6). Although changes were attenuated at month 6 (Figure 2) these remained significant for PAL (L0: +0.02, *P*=.04 and L1 to L3: +0.01, *P*=.006), sedentary time (L0: -199 min/wk, *P*<.001 and L1 to L3: -179 min/wk, *P*<.001), and MPA (L0: +27 min/wk, *P*=.002 and L1 to L3: +18 min/wk, *P*<.001).

However, there were no significant differences in objectively measured PA between individuals in personalized groups and matched controls (Table 2 and Figure 2).

At month 6, the change from baseline in MPA was significantly larger for L3 compared with L2 (L3: +32 min/wk vs L2: +7 min/wk, *P*=.04), but there were no other significant differences in PA between personalized groups (see Figure 2 and Multimedia Appendix 7). Results were unchanged when analyzing men and women separately (data not shown).

Effect of Different Levels of Personalized Advice on Self-Reported Physical Activity

Total Cohort: Generic Approach

Participants reported significant improvements in PA after the 6-month intervention (see Multimedia Appendix 8). Compared with the control group, individuals in personalized groups had significantly higher leisure time (nonsport) index (2.4%, *P*=.02) and total activity index scores (1.6%, *P*=.03) (see Multimedia Appendix 8). However, no significant differences were found between personalized groups (see Multimedia Appendix 8). Similar results were found at month 3 (see Multimedia Appendix 9).

Participants Who Received Advice to Increase Physical Activity and Matched Controls: Targeted Approach

After 6 months, there were significant improvements in self-reported PA during sport, leisure time (nonsport), and total PA among participants who received tailored advice (Figure 3). Compared with the control group, scores reported in personalized groups were significantly higher for leisure time (nonsport) index (3.6%, *P*=.009) and total activity index (2.5%, *P*=.009) (see Table 2 and Figure 3). Similar results were found at month 3 (see Multimedia Appendix 10). Finally, we also observed significant differences between personalized groups at month 6, scores for both indices being higher (3.9%, *P*=.006 and 2.9%, *P*=.008, respectively) for participants in L2 and L3 compared with L1 (see Figure 3 and Multimedia Appendix 7). Results were unchanged when analyzing men and women separately.

Importantly, results were also similar when including only individuals with both objective and self-reported PA (ie, completers, compliant with wearing the accelerometer and who have self-reported data) in the analysis (data not shown).

Figure 2. Changes from baseline to month 6 in physical activity measured objectively, for participants who received advice to increase physical activity (personalized groups Levels 1, 2, and 3) and matched controls (Level 0)—targeted approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, change in body weight, and change in accelerometer wear time. Individuals in Levels 1 (L1), 2 (L2), and 3 (L3) received personalized physical activity feedback based on current physical activity level (L1 to L3), phenotypic information (L2 and L3), and genotypic information (L3), whereas controls (L0) received nonpersonalized guidelines on physical activity.

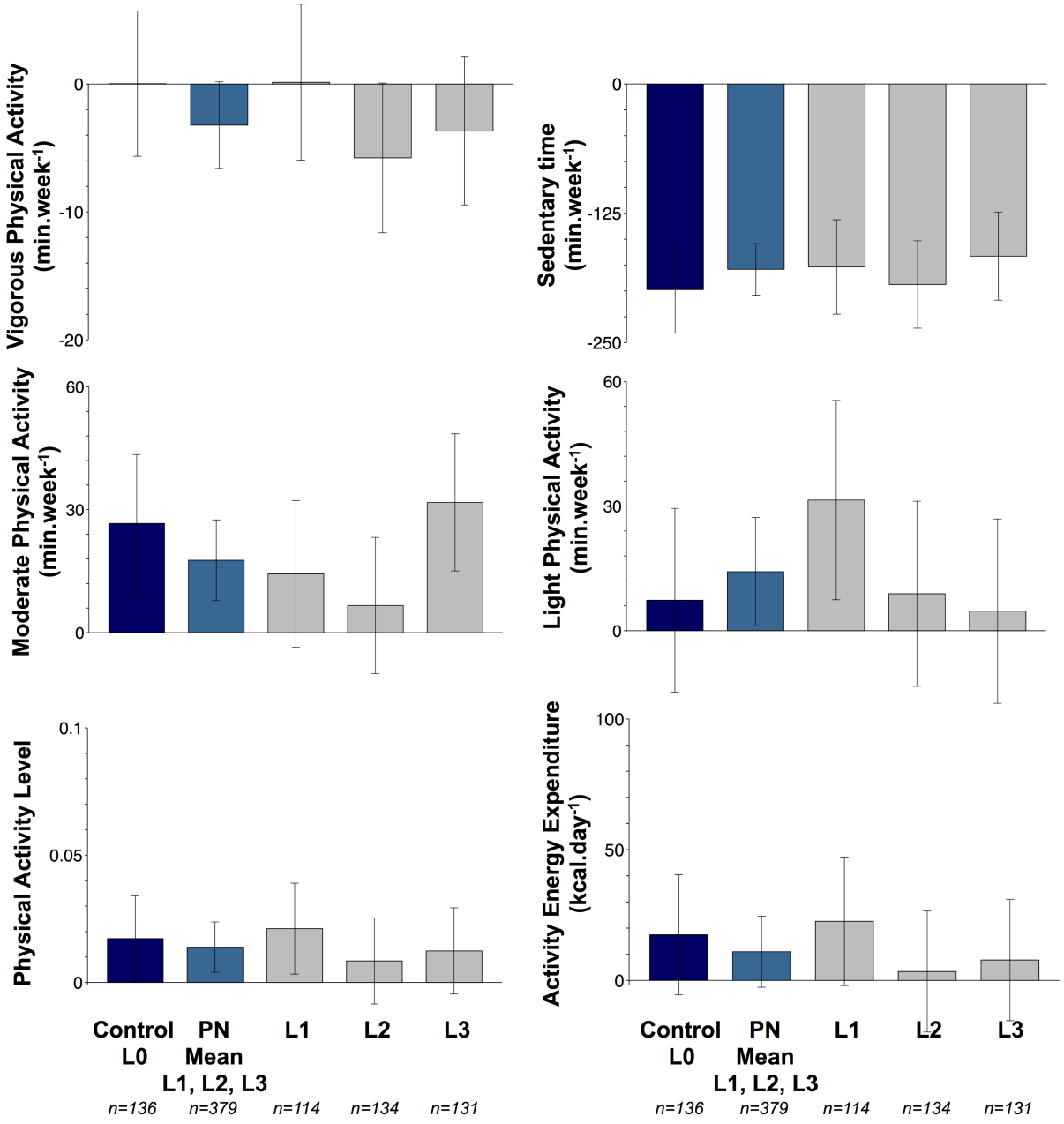


Table 2. Effect of targeted intervention on physical activity at month 6[a].

PA[b] components	Matched control, L0[c] (n=136), mean (SD)	Personalized advice, L1 to L3[d] (n=379), mean (SD)	Intervention effects, (L1 to L3) - L0 (95% CI)	*P*, L0 vs (L1 to L3)
Objective PA				
PAL[e]	1.68 (0.10)	1.68 (0.10)	-0.003 (-0.020 to 0.020)	.73
AEE[f] (kcal/d)	785 (137)	778 (135)	-6 (-34 to 21)	.64
Sedentary time (min/wk)	5182 (250)	5202 (247)	20 (-30 to 69)	.44
LPA[g] (min/wk)	479 (132)	486 (129)	7 (-19 to 33)	.59
MPA[h] (min/wk)	216 (100)	207 (98)	-9 (-29 to 11)	.36
VPA[i] (min/wk)	48 (34)	45 (34)	-3 (-10 to 3)	.33
Moderate-equivalent PA[j] (min/wk)	323 (154)	310 (154)	-14 (-45 to 17)	.35
Moderate-equivalent PA in bouts (min/wk)	140 (103)	131 (102)	-10 (-30 to 11)	.35
Self-reported PA				
Total activity index[k]	7.58 (0.87)	7.77 (0.87)	0.18 (0.05 to 0.32)	*.009*[l]
Work index[k]	2.24 (0.30)	2.27 (0.30)	0.03 (-0.02 to 0.07)	.26
Sport index[m]	2.58 (0.50)	2.65 (0.50)	0.070 (-0.005 to 0.150)	.07
Leisure time (nonsport) index[m]	2.77 (0.48)	2.87 (0.48)	0.10 (0.03 to 0.17)	*.009*

[a]Analysis is restricted to participants randomized to Levels 1 to 3 (L1, L2, and L3) who received personalized advice to increase PA, and to matched control group (L0) participants who would have received personalized advice to increase PA if they had been in a personalized group and not in L0. Data are presented as adjusted means and as the difference between the personalized groups (mean L1, L2, L3) and control with the corresponding 95% confidence interval. Differences between levels of personalized advice are presented in Multimedia Appendix 7. All analyses were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight. In addition, for objective PA variables analyses were adjusted for change in accelerometer wear time.

[b]Physical activity (PA).

[c]Level 0 (L0).

[d]Level 1 to Level 3 (L1 to L3).

[e]Physical activity level (PAL).

[f]Activity energy expenditure (AEE).

[g]Light-intensity PA (LPA).

[h]Moderate-intensity PA (MPA).

[i]Vigorous-intensity PA (VPA).

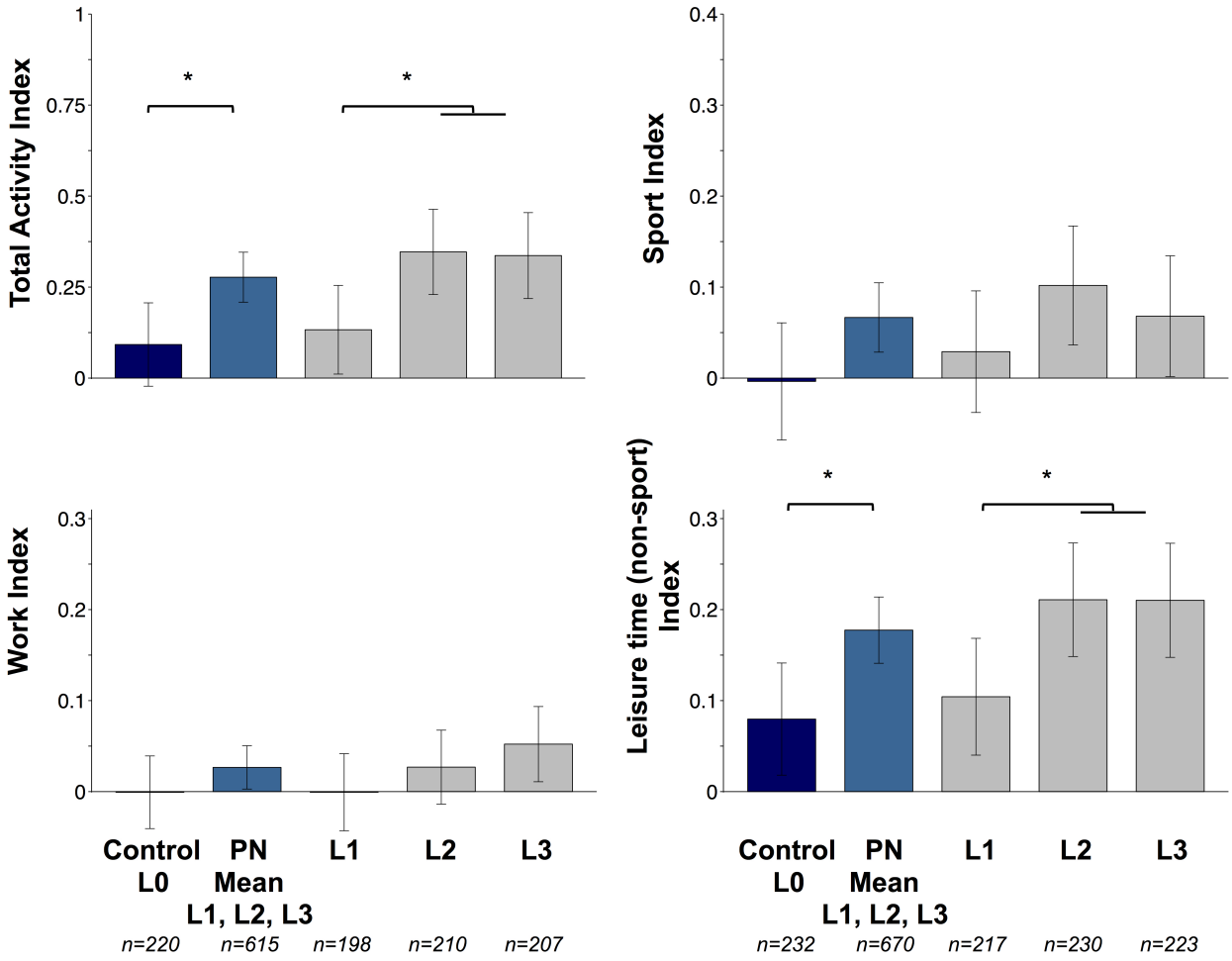
[j]Moderate-equivalent PA is (MPA + 2 × VPA).

[k]Participant numbers within each group are as follows: n=220 (L0), 198 (L1), 210 (L2), 207 (L3), and 615 (pooled L1, L2, and L3). For retired or unemployed individuals, work index, and therefore total index, cannot be calculated.

[l]Values in italics represent significant results.

[m]Participant numbers within each group are as follows: n=232 (L0), 217 (L1), 230 (L2), 223 (L3), and 670 (pooled L1, L2, and L3).

Figure 3. Changes from baseline to month 6 in self-reported physical activity (Baecke questionnaire) for participants who received advice to increase physical activity (personalized groups Levels 1, 2, and 3) and matched controls (Level 0)—targeted approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight. *Significant differences at P<.01. Individuals in Levels 1 (L1), 2 (L2), and 3 (L3) received personalized physical activity feedback based on current physical activity level (L1 to L3), phenotypic information (L2 and L3), and genotypic information (L3), whereas controls (L0) received nonpersonalized guidelines on physical activity.



Discussion

Principal Findings

For individuals who were not sufficiently active at baseline, as well as for the total cohort, personalized PA advice delivered via the Internet was more effective in improving self-reported PA compared with conventional "one size fits all" population-based advice. In addition, after 6 months of intervention, including phenotypic and/or genotypic information in the derivation of personalized PA advice led to bigger changes in self-reported PA than personalization based on diet and PA alone. However, these findings were not confirmed when PA was assessed objectively using accelerometers. Although we found some small significant improvements in objectively measured PA over the 6-month intervention, these changes were similar in all interventions groups.

Comparison With Previous Work

Most studies that have investigated the effectiveness of eHealth computer-tailored PA interventions have relied on self-reports [8], and their findings should be interpreted with caution. Few studies were identified that used objective PA outcome measures based on accelerometry [14-17,32]. Godino et al [15] noted that personalized PA feedback increased awareness of PA but did not promote change in PA after 2 months of follow-up. However, Hurling et al [32] found that participants who had access to a fully automated Internet, email, and mobile phone behavior change system—which included feedback on activity level and modules designed to help participants identify their perceived barriers and offered tailored solutions—had significantly higher objectively measured PA during the 9-week intervention compared with controls, who received verbal recommendations on PA but had no access to the behavior change system and received no feedback. Self-reported leisure PA was also significantly higher and time spent sitting significantly lower in the intervention group compared with the controls, but overall self-reported PA was similar in both groups [32]. In a study of older adults, Wijsman et al [17] observed that participants in an Internet program aiming to increase PA by monitoring and feedback by accelerometer and digital coaching had a significant increase of 11 minutes per day in moderate and vigorous PA after 3 months, whereas the

wait-listed controls showed no change in PA. Ashe et al [14] found significant increases in PA, as well as group differences at month 6, but their sample was small—13 participants in the intervention group and 12 in the control group—and group-based education and social support was included in addition to their Web-based intervention. Finally, Wanner et al [16] reported some improvements in self-reported PA after 6 weeks and 13 months of follow-up, but no differences between individuals in tailored and control groups, and no improvement in objectively measured PA for any group. These discrepancies between self-reported and objectively measured PA results are in line with our study. However, we found greater improvements in self-reported PA in tailored groups as compared with the controls. It could be that participants desired to comply with recommendations and that receiving more personalized feedback (Levels 2 and 3) increased this desire further in our study. It could also be that participants truly believed that they became more active when they actually did not. Contrary to Wanner et al, objective PA also improved, slightly but significantly, after 6 months of intervention, especially in participants who were inactive at baseline, in line with the results of Wijsman et al and Ashe et al. Yet in our study, those changes in objective PA were similar across all groups.

The number of studies testing the effectiveness of personalized feedback versus conventional population-based guidelines is limited. However, most authors stress the need for new ways to increase compliance and engagement of participants to ultimately successfully improve PA. Of those who started the Food4Me Study, 85.81% (1270/1480) completed the trial, indicating that our Web-based intervention was effective in retaining participants. However, only 49.32% (730/1480) of the starters had sufficient valid accelerometer data after 6 months. Thus, compliance with the study protocol in wearing the accelerometer remains a major issue, especially because those who were less compliant in our study had significantly lower baseline PAL.

Strengths and Limitations

The Food4Me Study is the largest Internet-based RCT to date to test the effects of personalized feedback on PA, using objectively measured PA with accelerometers. To our knowledge, it is also the first study to assess the effects of different levels of personalization including tailored phenotypic and genotypic information. Another strength is the inclusion of seven European centers that delivered the intervention with the same standardized protocol [18,33].

An important limitation is the relatively low compliance after 6 months with respect to accelerometer wear, which is lower than in other studies [16]. In our study, participants were asked to wear their accelerometer every day for the entire duration of the study, which may have been too demanding. Most participants did wear the accelerometer but not enough (ie, fewer than 3 valid weekdays and 2 valid weekend days). In other studies, participants received a PA monitor, were asked to wear it for the measurement period only (typically, 7 days), and to return the device immediately after to the researchers [15-17]. Better compliance in wearing the monitor may be obtained by having coaches motivating participants regularly [17], but this is not always feasible in a large-scale study. Advertisement for the study was primarily focused on personalized nutrition (ie, improving nutritional intakes) and not on PA. Moreover, the Food4Me Study was a multiple-behavior intervention including a large amount of information with extensive feedback, and many individuals may have felt they did not have time to try to make changes in both PA and diet concurrently [34].

A potential explanation as to why participants in personalized groups did not do better than controls could be that the TracmorD PA monitor used in the study constituted a basic, yet personalized, feedback by itself. That is, when the monitor is set on a flat surface at any time (eg, a table), lights on the monitor turn green depending on how much activity has been registered during the day. The more activity, the more lights turn green, for all participants including controls. Ideally, there should be no feedback on PA at all in the control group. This may explain the small but significant improvement in the control group. One could argue that providing feedback every 3 months might not be sufficiently frequent. In our study, half of the participants in the personalized groups received additional feedback based on measurements after 1 and 2 months (ie, high-intensity feedback: four feedback reports at months 0, 1, 2, and 3) but changes in objectively and self-reported PA at 3- and 6-month follow-up were similar to the other half of participants (ie, low-intensity feedback: feedback reports at months 0 and 3 only). Compliance in wearing the monitor did not differ between high- and low-intensity participants (data not shown). Furthermore, we cannot exclude the fact that phenotypic and genotypic characterization, and therefore feedback to participants in Levels 2 and 3, may not have been optimal for PA-related outcomes. For example, our only PA-related genotypic variant was in the *FTO* gene and perhaps just one gene variant would be insufficient to motivate participants in Level 3 to increase PA beyond those in Level 2.

Finally, although accelerometry is an objective measure of PA, it can underestimate certain activities, such as carrying heavy loads and when the torso remains relatively static (eg, during cycling). Accelerometry cannot (easily) distinguish PA when ascending (eg, walking uphill) from movement on the flat yet there could be large differences in energy expenditure between the two types of movement. Our monitor was waterproof and could be worn during swimming, but underestimation of activity intensity is common. Nevertheless, the TracmorD has been validated against the doubly labeled water method and several publications show that it is a reliable and accurate monitor [12,13,20,35]. Although devices may not capture all types of PA, questionnaires have been shown repeatedly to be inaccurate, often overestimating PA [11,36]. The Baecke questionnaire, although extensively validated [28,29], is no exception [37]. Our results support the position adopted by others that self-reported measures of PA should be interpreted with caution and preferably not be used to draw conclusions on the effectiveness of PA interventions [8,38]. Thus, it is better for personalized feedback to have objective measures of PA such as accelerometry. Such technologies are becoming very relevant tools for both surveys and interventions to promote public health. They are developing rapidly and are commonly available for download to mobile phones and watches, allowing greater

accessibility for the general population, which may help with noncompliance. However, these new apps will need to be rigorously tested.

Conclusions

We observed small but significant improvements in objectively measured PA after 3 and 6 months of intervention, although changes were similar across all groups. Personalized advice on PA did not promote larger and more sustained improvements in objective PA as compared with a conventional "one size fits all" approach delivered via the Internet. Furthermore, increasing the degree of personalization using phenotypic or genotypic information had no effect on changes in objectively measured PA compared with personalized feedback based on diet and PA alone. Based on self-reports, however, PA improved significantly more with higher degrees of personalized advice.

Acknowledgments

The Food4Me Study was supported by the European Commission under the Food, Agriculture, Fisheries and Biotechnology Theme of the 7th Framework Programme for Research and Technological Development (grant number 265494).

Authors' Contributions

JCM was the Food4Me intervention study coordinator. CFMM and WHMS drafted the paper. CFMM and CCM performed the statistical analysis for the manuscript. CAD, ERG, LB, YM, IT, JAL, JAM, WHMS, HD, MG, and JCM contributed to the research design. CCM, CFMM, CW, HF, CBO, RF, ALM, JAL, SNC, RSC, SK, AS, MG, CPL, GM, MCW, JH, and JCM contributed to developing the Standardized Operating Procedures for the study. CFMM, CCM, SNC, RSC, CW, HF, COB, RF, ALM, SK, CPL, GM, AS, MG, MCW, and JCM conducted the intervention. CFMM, WHMS, CCM, AG, and JH contributed to physical activity measurements. All authors contributed to a critical review of the manuscript. All authors approved the final version to be published.

Conflicts of Interest

AG and JH are employed by Philips. Others have no conflicts of interest to declare.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.1 [39].

[PDF File (Adobe PDF File), 692KB - jmir_v17i10e231_app1.pdf]

Multimedia Appendix 2

Physical activity in feedback reports and on the Food4Me website.

[PDF File (Adobe PDF File), 1MB - jmir_v17i10e231_app2.pdf]

Multimedia Appendix 3

Baseline characteristics of the Food4Me participants who received personalized advice to increase physical activity and matched controls—targeted approach.

[PDF File (Adobe PDF File), 34KB - jmir_v17i10e231_app3.pdf]

Multimedia Appendix 4

Changes from baseline to month 3 in objectively measured physical activity for the total cohort—generic approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, change in body weight, and change in accelerometer wear time.

[PNG File, 802KB - jmir_v17i10e231_app4.png]

Multimedia Appendix 5

Changes from baseline to month 6 in objectively measured physical activity for the total cohort—generic approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, change in body weight, and change in accelerometer wear time.

[PNG File, 753KB - jmir_v17i10e231_app5.png]

Multimedia Appendix 6

Changes from baseline to month 3 in objectively measured physical activity for participants who received advice to increase PA (L1 to L3) and matched controls (L0)—targeted approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, change in body weight, and change in accelerometer wear time.

[PNG File, 825KB - jmir_v17i10e231_app6.png]

Multimedia Appendix 7

Effect of targeted intervention on physical activity for personalized groups at month 6.

[PDF File (Adobe PDF File), 39KB - jmir_v17i10e231_app7.pdf]

Multimedia Appendix 8

Changes from baseline to month 6 in self-reported physical activity (Baecke questionnaire) for the total cohort—generic approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight.

[PNG File, 547KB - jmir_v17i10e231_app8.png]

Multimedia Appendix 9

Changes from baseline to month 3 in self-reported physical activity (Baecke questionnaire) for the total cohort—generic approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight.

[PNG File, 544KB - jmir_v17i10e231_app9.png]

Multimedia Appendix 10

Changes from baseline to month 3 in self-reported physical activity (Baecke questionnaire) for participants who received advice to increase physical activity (L1 to L3) and matched controls (L0)—targeted approach. Data are presented as adjusted changes from baseline. Error bars represent 95% confidence intervals. Models were adjusted for baseline values, sex, age, country, smoking, baseline BMI, baseline season, and change in body weight.

[PNG File, 544KB - jmir_v17i10e231_app10.png]

References

1. Alwan A, Armstrong T, Cowan M, Riley L. Noncommunicable Diseases: Country Profiles 2011. Geneva, Switzerland: World Health Organization; 2011. URL: http://apps.who.int/iris/bitstream/10665/44704/1/9789241502283_eng.pdf [accessed 2015-10-07] [WebCite Cache ID 6c5zqLZpj]
2. Ekelund U, Ward H, Norat T, Luan J, May A, Weiderpass E, et al. Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). Am J Clin Nutr Mar 2015;101(3):613-621. [doi: 10.3945/ajcn.114.100065] [Medline: 25733647]
3. Hallal P, Andersen L, Bull F, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
4. Ronda G, Van Assema P, Brug J. Stages of change, psychological factors and awareness of physical activity levels in The Netherlands. Health Promot Int 2001 Dec;16(4):305-314 [FREE Full text] [Medline: 11733449]
5. Watkinson C, van Sluijs EM, Sutton S, Hardeman W, Corder K, Griffin S. Overestimation of physical activity level is associated with lower BMI: a cross-sectional analysis. Int J Behav Nutr Phys Act 2010;7:68 [FREE Full text] [doi: 10.1186/1479-5868-7-68] [Medline: 20854659]
6. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
7. Davies C, Spence J, Vandelanotte C, Caperchione C, Mummery W. 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]

8. Laplante C, Peng W. A systematic review of e-health interventions for physical activity: an analysis of study design, intervention characteristics, and outcomes. Telemed J E Health 2011 Sep;17(7):509-523. [doi: 10.1089/tmj.2011.0013] [Medline: 21718092]
9. 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]
10. Tehard B, Saris WH, Astrup A, Martinez J, Taylor M, Barbe P, et al. Comparison of two physical activity questionnaires in obese subjects: the NUGENOB study. Med Sci Sports Exerc 2005 Sep;37(9):1535-1541. [Medline: 16177606]
11. Celis-Morales C, Perez-Bravo F, Ibañez L, Salas C, Bailey ME, Gill JM. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. PLoS One 2012;7(5):e36345 [FREE Full text] [doi: 10.1371/journal.pone.0036345] [Medline: 22590532]
12. Dannecker K, Sazonova N, Melanson E, Sazonov E, Browning R. A comparison of energy expenditure estimation of several physical activity monitors. Med Sci Sports Exerc 2013 Nov;45(11):2105-2112 [FREE Full text] [doi: 10.1249/MSS.0b013e318299d2eb] [Medline: 23669877]
13. Plasqui G, Bonomi A, Westerterp K. Daily physical activity assessment with accelerometers: new insights and validation studies. Obes Rev 2013 Jun;14(6):451-462. [doi: 10.1111/obr.12021] [Medline: 23398786]
14. Ashe M, Winters M, Hoppmann C, Dawes M, Gardiner P, Giangregorio L, et al. "Not just another walking program": Everyday Activity Supports You (EASY) model—a randomized pilot study for a parallel randomized controlled trial. Pilot Feasibility Studies 2015 Jan 12;1(4):1-12. [doi: 10.1186/2055-5784-1-4]
15. Godino J, Watkinson C, Corder K, Marteau T, Sutton S, Sharp S, et al. Impact of personalised feedback about physical activity on change in objectively measured physical activity (the FAB study): a randomised controlled trial. PLoS One 2013;8(9):e75398 [FREE Full text] [doi: 10.1371/journal.pone.0075398] [Medline: 24066178]
16. Wanner M, Martin-Diener E, Braun-Fahrländer C, Bauer G, Martin B. Effectiveness of active-online, an individually tailored physical activity intervention, in a real-life setting: randomized controlled trial. J Med Internet Res 2009;11(3):e23 [FREE Full text] [doi: 10.2196/jmir.1179] [Medline: 19666456]
17. Wijsman C, Westendorp R, Verhagen E, Catt M, Slagboom P, 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]
18. Celis-Morales C, Livingstone K, Marsaux CF, Forster H, O'Donovan C, Woolhead C, et al. Design and baseline characteristics of the Food4Me study: a Web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr 2015 Jan;10(1):450 [FREE Full text] [doi: 10.1007/s12263-014-0450-2] [Medline: 25491748]
19. Food4Me. 2011. URL: http://www.food4me.org [WebCite Cache ID 6XJSSOSJD]
20. Bonomi A, Plasqui G, Goris AH, Westerterp K. 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]
21. Philips DirectLife. URL: http://www.directlife.philips.com [WebCite Cache ID 6XJUgQaGy]
22. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
23. R Core Team. The R Project. Vienna, Austria: The R Foundation The R Project for Statistical Computing URL: https://www.r-project.org/ [accessed 2015-10-07] [WebCite Cache ID 6c60fowbl]
24. Henry CJ. Basal metabolic rate studies in humans: measurement and development of new equations. Public Health Nutr 2005 Oct;8(7A):1133-1152. [Medline: 16277825]
25. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva, Switzerland: World Health Organization; 2010. Recommended population levels of physical activity for health URL: http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf [accessed 2015-10-08] [WebCite Cache ID 6c80nAiJU]
26. Troiano R, Berrigan D, Dodd K, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
27. Baecke J, Burema J, Frijters J. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982 Nov;36(5):936-942 [FREE Full text] [Medline: 7137077]
28. Philippaerts RM, Westerterp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. Int J Sports Med 1999 Jul;20(5):284-289. [doi: 10.1055/s-2007-971132] [Medline: 10452224]
29. Montoye HJ, Kemper HCG, Saris WHM, Washburn RA. In: Montoye HJ, editor. Measuring Physical Activity and Energy Expenditure. Champaign, IL: Human Kinetics; 1996.
30. Celis-Morales C, Livingstone KM, Woolhead C, Forster H, O'Donovan CB, Macready AL, et al. How reliable is Internet-based self-reported identity, socio-demographic and obesity measures in European adults? Genes Nutr 2015 Sep;10(5):476 [FREE Full text] [doi: 10.1007/s12263-015-0476-0] [Medline: 26143178]
31. Rousseeuw P, Croux C, Todorov V, Ruckstuhl A, Salibian-Barrera M, Verbeke T, et al. The Comprehensive R Archive Network, R Project. Vienna, Austria: The R Foundation; 2015. robustbase: basic robust statistics. R package version 0.92-5 URL: http://cran.r-project.org/web/packages/robustbase/index.html [accessed 2015-10-07] [WebCite Cache ID 6c62dJAIi]

32. 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]
33. Livingstone K, Celis-Morales C, Navas-Carretero S, San-Cristobal R, O'Donovan C, Forster H, et al. Profile of European adults interested in Internet-based personalised nutrition: the Food4Me study. Eur J Nutr 2015 Apr 17:1-11. [doi: 10.1007/s00394-015-0897-y] [Medline: 25893715]
34. Norman G, Zabinski M, Adams M, Rosenberg D, Yaroch A, Atienza A. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345. [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
35. Valenti G, Camps SG, Verhoef SP, Bonomi AG, Westerterp KR. Validating measures of free-living physical activity in overweight and obese subjects using an accelerometer. Int J Obes (Lond) 2014 Jul;38(7):1011-1014. [doi: 10.1038/ijo.2013.195] [Medline: 24166066]
36. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. J Phys Act Health 2010 Jul;7(4):541-550. [Medline: 20683097]
37. Walsh MC, Hunter GR, Sirikul B, Gower BA. Comparison of self-reported with objectively assessed energy expenditure in black and white women before and after weight loss. Am J Clin Nutr 2004 Jun;79(6):1013-1019 [FREE Full text] [Medline: 15159231]
38. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TI, Energy Balance Measurement Working Group. Energy balance measurement: when something is not better than nothing. Int J Obes (Lond) 2015 Jul;39(7):1109-1113 [FREE Full text] [doi: 10.1038/ijo.2014.199] [Medline: 25394308]
39. Eysenbach G. CONSORT-EHEALTH: Improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

AEE: activity energy expenditure
BMI: body mass index
BMR: basal metabolic rate
CIBERobn: Centro de Investigación Biomédica en Red-Fisiopatología de la Obesidad y Nutrición
FTO: fat mass and obesity associated
GP: general practitioner
IPAQ: International Physical Activity Questionnaire
IZZ: National Food & Nutrition Institute
KASP: competitive allele-specific polymerase chain reaction
L0: Level 0
L1: Level 1
L2: Level 2
L3: Level 3
LPA: light-intensity physical activity
MET: metabolic equivalent
MPA: moderate-intensity physical activity
MUMC+: Maastricht University Medical Centre +
NUTRIM: (School of) Nutrition and Translational Research in Metabolism
OR: odds ratio
PA: physical activity
PAL: physical activity level
RCT: randomized controlled trial
SNP: single nucleotide polymorphism
UCD: University College Dublin
VPA: vigorous-intensity physical activity
WHO: World Health Organization
ZIEL: Zentralinstitut für Ernährungs- und Lebensmittelforschung

Edited by P Bamidis; submitted 14.05.15; peer-reviewed by Y Lagerros, A Billis; comments to author 29.06.15; revised version received 20.07.15; accepted 22.09.15; published 14.10.15.

Please cite as:
Marsaux CFM, Celis-Morales C, Fallaize R, Macready AL, Kolossa S, Woolhead C, O'Donovan CB, Forster H, Navas-Carretero S, San-Cristobal R, Lambrinou CP, Moschonis G, Surwillo A, Godlewska M, Goris A, Hoonhout J, Drevon CA, Manios Y, Traczyk I, Walsh MC, Gibney ER, Brennan L, Martinez JA, Lovegrove JA, Gibney MJ, Daniel H, Mathers JC, Saris WHM
Effects of a Web-Based Personalized Intervention on Physical Activity in European Adults: A Randomized Controlled Trial
J Med Internet Res 2015;17(10):e231
URL: http://www.jmir.org/2015/10/e231/
doi:10.2196/jmir.4660
PMID:26467573

Original Paper

Comparing a Video and Text Version of a Web-Based Computer-Tailored Intervention for Obesity Prevention: A Randomized Controlled Trial

Michel Jean Louis Walthouwer[1,2], MSc; Anke Oenema[1,2], PhD; Lilian Lechner[2,3], PhD; Hein de Vries[1,2], PhD

[1]Maastricht University, Department of Health Promotion, Maastricht, Netherlands

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

[3]Open University, Heerlen, Netherlands

Corresponding Author:
Michel Jean Louis Walthouwer, MSc
Maastricht University
Department of Health Promotion
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 43 3882402
Fax: 31 43 3882402
Email: michel.walthouwer@maastrichtuniversity.nl

Abstract

Background: Web-based computer-tailored interventions often suffer from small effect sizes and high drop-out rates, particularly among people with a low level of education. Using videos as a delivery format can possibly improve the effects and attractiveness of these interventions

Objective: The main aim of this study was to examine the effects of a video and text version of a Web-based computer-tailored obesity prevention intervention on dietary intake, physical activity, and body mass index (BMI) among Dutch adults. A second study aim was to examine differences in appreciation between the video and text version. The final study aim was to examine possible differences in intervention effects and appreciation per educational level.

Methods: A three-armed randomized controlled trial was conducted with a baseline and 6 months follow-up measurement. The intervention consisted of six sessions, lasting about 15 minutes each. In the video version, the core tailored information was provided by means of videos. In the text version, the same tailored information was provided in text format. Outcome variables were self-reported and included BMI, physical activity, energy intake, and appreciation of the intervention. Multiple imputation was used to replace missing values. The effect analyses were carried out with multiple linear regression analyses and adjusted for confounders. The process evaluation data were analyzed with independent samples *t* tests.

Results: The baseline questionnaire was completed by 1419 participants and the 6 months follow-up measurement by 1015 participants (71.53%). No significant interaction effects of educational level were found on any of the outcome variables. Compared to the control condition, the video version resulted in lower BMI (B=-0.25, *P*=.049) and lower average daily energy intake from energy-dense food products (B=-175.58, *P*<.001), while the text version had an effect only on energy intake (B=-163.05, *P*=.001). No effects on physical activity were found. Moreover, the video version was rated significantly better than the text version on feelings of relatedness (*P*=.041), usefulness (*P*=.047), and grade given to the intervention (*P*=.018).

Conclusions: The video version of the Web-based computer-tailored obesity prevention intervention was the most effective intervention and most appreciated. Future research needs to examine if the effects are maintained in the long term and how the intervention can be optimized.

ClinicalTrial: Netherlands Trial Register: NTR3501; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3501 (Archived by WebCite at http://www.webcitation.org/6cBKIMaW1)

(J Med Internet Res 2015;17(10):e236) doi:10.2196/jmir.4083

KEYWORDS

randomized controlled trial; web-based; computer-tailoring; obesity; educational level; delivery strategy

Introduction

Overweight and obesity rates have increased rapidly during the last 30 years [1,2]. In 2008, around 900 million adults were overweight and 502 million were obese [3,4]. In many Western countries, these figures are significantly higher among people with a low level of education [5-9]. For example, in the Netherlands the prevalence of overweight is 64.4% among adults with a low level of education compared with 40.1% among adults with a high level of education [10].

Because overweight and obesity affect large numbers of people, these interventions should have the possibility to reach many people in an efficacious yet cost-effective manner [11]. Web-based computer-tailored interventions meet this requirement. These interventions aim to change people's health behavior by providing individually adapted information via the Internet [12]. Hence, they can be disseminated easily among a large target population for relatively low costs [11]. Research has already shown that Web-based computer-tailored interventions can have positive effects on physical activity, dietary intake, and body weight [11-16] and that they can be cost-effective [17,18]. Yet, the current evidence for the effectiveness for these interventions is inconclusive as effects are mostly small and are found only in the short term [19,20]. Moreover, Web-based computer-tailored interventions often suffer from high dropout rates that reach up to around 50% [21-24]. These interventions in particular have problems reaching people with a low educational level—the people most in need of change [25]. Hence, to optimize the potential of Web-based computer-tailored obesity prevention interventions, it is necessary to examine how their impact and attractiveness can be improved [25-27].

One possible solution may be to provide the information within these interventions by means of a delivery format that better fits the receivers' preferences [28,29]. Nearly all previous Web-based computer-tailored interventions have primarily used text-driven messages to provide information. However, particularly people with a low educational level generally are less text oriented [30]. Recent studies provide indications that the delivery of intervention content via videos may improve the effectiveness of Web-based computer-tailored interventions [26,31]. Although the current evidence for this hypothesis is not compelling, video messages could be more appropriate because people with a low educational level typically have more difficulties processing large amounts of text [32]. Videos may work better because they reduce the cognitive effort needed to process information, which can lead to better comprehension [33].

To examine whether the use of videos can indeed improve the effectiveness and attractiveness, we developed 2 versions of a Web-based computer-tailored intervention. This intervention aimed to achieve small changes in dietary intake and physical activity in order to prevent weight gain among Dutch adults with a healthy weight or with overweight, specifically, a body mass index (BMI) between 18.5 and 30 kg/m^2. Both versions of the intervention had exactly the same content but had a different information delivery format. One version was fully text based, without the use of visual elements (text version), and the other provided the core tailored information by means of videos (video version).

The main aim of this study was to examine the effects of the video and text version in comparison to a waiting list control condition on dietary intake, physical activity, and BMI among Dutch adults at 6 months' follow-up. A second study aim was to examine potential differences in participants' appreciation of the intervention between the video and text version. The final study aim was to examine possible differences in efficacy and appreciation per educational level. We hypothesized that the video version would be more effective and better appreciated, particularly among people with a low level of education.

Methods

The Ethical Committee of the Open University Heerlen reviewed the study protocol and had no objections. The study is registered in the Dutch Trial Register (NTR3501). See Multimedia Appendix 1 for the CONSORT EHEALTH checklist [34].

Study Design and Respondents

A three-armed randomized controlled trial was conducted with 2 experimental conditions (video and text intervention) and a waiting list control group that had the opportunity to use one of the interventions after the study. Measurements took place at baseline (T0) and 6 months (T1) after baseline. Criteria for participation were being at least 18 years old, having a paid job (because of initial recruitment procedure), a BMI between 18.5 and 30 kg/m^2, and sufficient command of the Dutch language. People with a physical condition that severely influenced their dietary or physical activity pattern (eg, diabetes) were not eligible to participate.

It was estimated that 2000 participants were needed to complete the baseline questionnaire in order to be able to detect a medium-sized effect (d=0.5) on BMI and behavior with a power of .90, a significance level of .05, and taking into account a dropout percentage of 50% between baseline and follow-up. This number of participants would also allow testing interaction effects between participants with a low, medium, and high level of education [26].

Procedure

Participants were recruited from September 2012 until February 2013. Participants were recruited during medical screenings by various occupational health centers, directly through companies, and via advertisements in national and local newspapers. All recruitment materials (ie, brochures, emails, advertisements) included information about the intervention study as well as a hyperlink to the study website where participants could register to participate. After registration and giving online informed consent, participants were randomly assigned to one of the 3

study conditions (ie, video version, text version, and control group) in a computer-determined sequence. After randomization, participants received a username and password by email. Participants were unaware of which study condition they were allocated to until they accessed the baseline questionnaire (T0). Two weeks after completion of this questionnaire, participants in the intervention conditions were given access to the intervention. Participants could use the assigned intervention for a maximum period of 3 months. Six months after baseline, participants were asked by email to fill out the online follow-up questionnaire (T1). To decrease the likelihood of attrition, participants were informed that they could win one of hundred cash prizes of €100 if they completed all questionnaires [35].

Intervention

The Web-based computer-tailored intervention was developed systematically using the Intervention Mapping protocol [36]. Detailed information about the development process and the content of the intervention can be found elsewhere [26]. The objective of the intervention was to prevent weight gain or achieve modest weight loss by making small changes in dietary intake and/or physical activity. In the video version, about 75% of the educational content was delivered via videos. The remaining 25% consisted of text-based content to give instructions about setting goals and making action and coping plans as well as for the delivery of optional in-depth information. The videos had a news-driven format in which professional actors read aloud the tailored information. This information was exactly the same as the information that could be read in text in the text version of the intervention. In both the video and text versions, the tailored information was based on participants' answers to online questions about their dietary intake, physical activity level, and sociocognitive beliefs (eg, self-efficacy). The feedback was very specific and, for example, clearly indicated which specific behavior changes participants could make (eg, decrease intake of chocolate with X per day).

The theoretical framework of the intervention consisted of a combination of self-regulation theories [37,38] and the I-Change Model [39]. Self-regulation theories were in particular used as input for the general framework for the intervention. Accordingly, the intervention aimed to create awareness of behavior, identify areas for change, set goals and make plans, and finally start and monitor the behavior change. The I-Change Model has mainly been used to make people aware of their behavior and for indicating which behavior change participants were most motivated to do. In these sessions, participants received feedback about their motivational beliefs (eg, attitude and self-efficacy) and could make action plans. In line with these theories, the following behavior change methods were used: consciousness raising, tailored feedback on behavior and cognitions, goal setting, action and coping planning, and evaluation of goal pursuit.

The intervention consisted of 6 weekly sessions, and each session lasted about 15 minutes. After Session 1, participants could continue to Session 2 directly. Hence, between Sessions 1 and 2 there was no mandatory waiting period (in contrast to the subsequent sessions). Figures 1 and 2 provide an example of the video and text versions, respectively.

Figure 1. Example of the video version of the intervention.

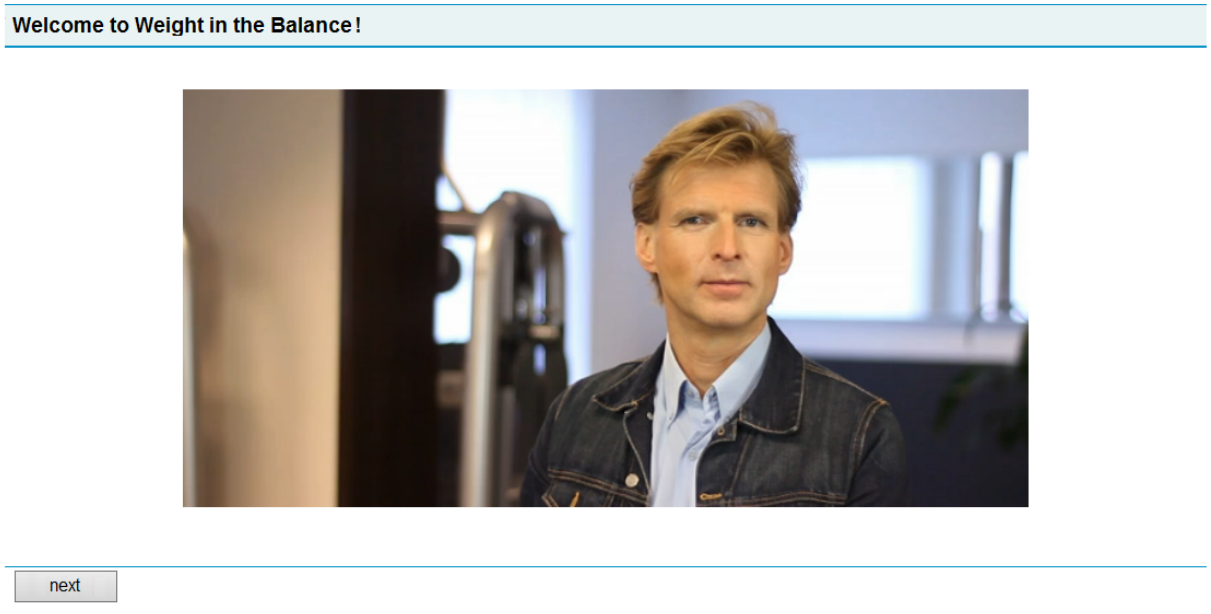
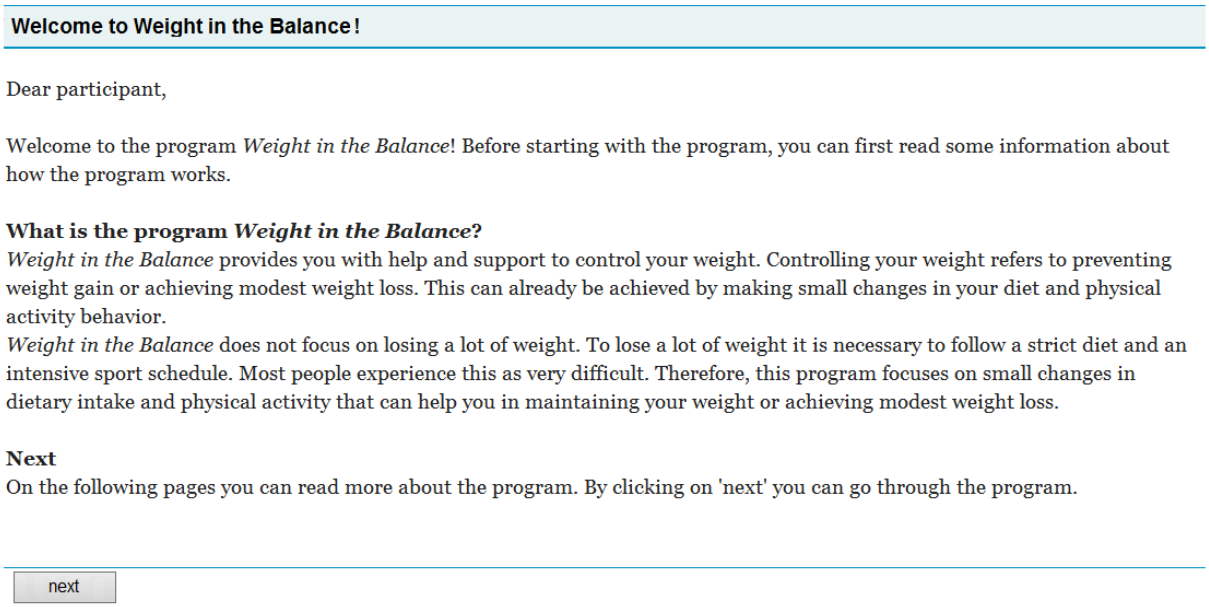


Figure 2. Example of the text version of the intervention.



Session 1

The aim of Session 1 was to inform participants about the different intervention sessions. Next, participants were provided with tailored feedback about their weight, behavior (dietary intake and physical activity), and sociocognitive beliefs toward improving their diet and physical activity level (risk perception, attitude, self-efficacy, and social influence). The aim of this feedback was to indicate which changes would best fit the participant. After receiving this information, participants subsequently had to set a goal by deciding if they wanted to maintain their current weight or lose a little weight. Participants also had to decide if they wanted to improve their physical activity level, their dietary intake, or both. To help participants with setting these 2 goals, they received information about the purpose of setting goals and examples of adequate goals.

Session 2

The aim of Session 2 was to provide participants with detailed feedback on the chosen behavior in order to inform them which small changes they could make to achieve their weight goal. Based on this information, participants could make “if then” plans by specifying when, where, and how they were going to undertake the behavior change. To further help participants with this, they received instructions about how to make appropriate plans as well as examples of good plans. After Session 2, participants could start realizing their goals and plans.

Session 3

The aim of Session 3 was to help participants carry out and maintain the behavior change. For this purpose, participants first received tailored feedback about their behavior change progress. This feedback was given by assessing participants’ current behavior and comparing this to their weight and behavior reported in Session 1. Based on this comparison, it was indicated whether or not participants’ behavior had improved and if they had achieved their goal. In addition, participants were also given the option to make coping plans. For this purpose, participants first received information about the purpose of coping planning. Next, participants could indicate which difficult situations they had encountered. For each selected situation (eg, being hungry), participants received tips about how to deal with this situation (eg, eat something with fewer calories such as fruit). Based on this feedback, participants could eventually make their own coping plan by selecting their own preferred coping response from a list with predefined options.

Sessions 4-6

The last 3 sessions were identical to the third session, but each new session also consisted of 1 or 2 new elements. Session 4, for example, also consisted of narratives in which a role model told how their behavior change was going and how they dealt with difficult situations. Participants were also given the possibility to change their goals and action plans. Session 5 was similar to Session 4, but in this session, participants received tailored feedback for the first time on their weight change by indicating whether or not they had achieved their weight goal. Finally, Session 6 was again similar to the previous session but additionally addressed the topic of how to maintain behavior changes in the long term. For this purpose, participants had the possibility of setting a long-term weight goal and making an action plan for achieving this goal. This last session ended with a review of the essential elements of the whole intervention.

Measurements

Outcome Variables

All outcome variables (ie, BMI, dietary intake, and physical activity) were assessed using online self-reports at both T0 and T1. Participants who had not completed the online follow-up questionnaire (T1) after several email reminders were contacted by telephone to assess their body weight.

First, participants’ body weight in kilograms and height in meters were assessed in order to calculate their BMI. To improve the adequacy of reporting, participants were asked to indicate their weight in the morning without clothes and shoes.

Dietary intake was assessed by means of a food frequency questionnaire consisting of 66 items, which was based on a validated questionnaire to assess fat intake [40]. The intake levels of mainly energy-dense products from 6 different food categories were assessed (ie, dairy products, sandwiches and fillings, food at dinner, sweet and savory snacks, hot and cold beverages, and alcohol). For each food product, the frequency (ie, number of days per week) and quantity (ie, servings per day) were assessed. When applicable, type of product (eg, use of skimmed, semi-skimmed, or whole milk) and portion size (eg, size of candy bar) were assessed as well. For each food product, the average daily intake was calculated. This was subsequently combined with the energy value of each food product [41] in order to calculate a score for the average daily intake of calories from energy-dense food products.

Physical activity was assessed using the Short Questionnaire to Assess Health-Enhancing Physical Activity (SQUASH) [42]. This questionnaire has proven to be a reliable and valid tool to estimate the level of physical activity among Dutch adults [43]. The SQUASH assesses participants' level of physical activity per category (ie, commuting activities, leisure time activities, household activities, and activities at work). For each activity, participants had to indicate how many days per week they engaged in this activity, average time per day spent in doing this activity, and the intensity of the activity (light, moderate, or vigorous). Based on these questions, a total score was calculated for the average daily minutes of moderate-to-vigorous intensity physical activity.

Demographics

All demographics were assessed at T0. Demographic variables consisted of gender, age, and educational level (ie, the highest level of education completed), which was categorized into low (primary or basic vocational school), medium (secondary vocational school or high school), and high (higher vocational school or university) [44].

Sociocognitive Variables

All sociocognitive variables (ie, self-efficacy, intention, and self-regulation skills) were assessed at T0. For this purpose, adapted measures of previous studies [27,45,46] were used, including a 5-point Likert answering scale ranging from 1 (low) to 5 (high). A scale was computed by calculating a mean score.

Participants' self-efficacy was measured separately for physical activity (alpha=.83) and dietary intake (alpha=.81) using 4 items per behavior. Participants were asked, for example, about their confidence and ability to improve their diet and physical activity level.

Intention was measured with 1 item per behavioral outcome by asking participants if they intended to improve their diet and physical activity level within the next 6 months.

Self-regulation skills were measured for the types of skills that are important for successfully translating intentions into behavior change (ie, goal setting, action planning, monitoring, and coping planning). Items were derived from existing instruments [47,48]. Goal setting (alpha=.72) was measured with 3 items by asking participants if they set a goal in advance when, for example, they want to manage their weight. Next, action planning was measured with 3 items per behavioral outcome. Participants were asked if they had a clear plan when, where, and how they wanted to improve their diet (alpha=.90) and physical activity level (alpha=.94). Monitoring (alpha=.74) was measured using 4 items that assessed to which degree participants monitored their weight and behavior on a regular basis. Finally, 2 items per behavioral outcome were used to assess coping planning. Participants were asked to which degree they were able to identify hindering situations in advance and thought that they were able to deal with these situations for both dietary intake (alpha=.70) and physical activity (alpha=.72).

Process Evaluation

Appreciation of the intervention was assessed at T1 using a 5-point Likert scale ranging from 1 (low) to 5 (high). Using 1 item per variable, participants were first asked to indicate to which degree they thought the information and feedback in the intervention was interesting, useful, understandable, and fitted to their own situation. Participants were also asked to give an overall rating of the intervention on a scale ranging from 1 (low) to 10 (high). Last, participants were asked about their feelings of autonomy, relatedness, and competence during the intervention. These concepts were derived from Self-Determination Theory [49], and the items were developed using existing questionnaires [21,27,45]. For these 3 concepts, average scale scores were computed. Autonomy (alpha=.88) was assessed by 2 items. Participants were asked if they had the feeling that they could decide by themselves which goals they could set and which information they could read in the intervention. Relatedness (alpha=.92) was assessed with 3 items by asking participants if they felt involved and supported by the intervention. Competence (alpha=.93) was assessed with 3 items by asking participants if the intervention had increased their confidence in their ability to manage their weight, dietary intake, and physical activity behavior. Finally, login data was used to assess use of the intervention.

Statistical Analyses

At both T0 and T1, multiple imputation was used to replace missing values [50,51]. Descriptive statistics and frequencies were used to describe the characteristics of the study population and the overall flow through the study. Baseline differences between the 3 study conditions were examined using analyses of variance with Tukey post hoc tests for continuous variables and chi-square tests with Bonferroni correction (P=.05/P=.017) for categorical variables. To examine the possible presence of selective attrition between baseline and follow-up, a logistic regression analysis was performed with attrition at follow-up as outcome (completed T1=0, not completed T1=1) and study condition and all baseline variables as predictors.

The effect analyses were conducted for each outcome variable separately (BMI, dietary intake, physical activity) using linear regression analyses with the enter method. The effects of the intervention conditions were compared to the control condition for which the study condition variable was recoded into 2 dummies (ie, video versus control and text versus control). The analyses were adjusted for potential confounders (ie, baseline behavior, predictors of attrition, and baseline differences) by

including these variables as covariates. The analyses also included study condition × educational level interaction terms to assess potential educational differences in intervention effects. Cohen's *d* effect sizes were calculated for all outcome variables [52]. As secondary analyses, we also compared the effects of the intervention conditions with each other. Moreover, the analyses were performed with both a complete case and multiple imputation dataset.

Finally, the process evaluation data were analyzed using linear regression analyses with the enter method. These analyses included study condition × educational level interaction terms to identify potential educational differences in appreciation. When no interaction effects were found, independent samples *t* tests were conducted to examine differences between the video and text conditions on the process evaluation variables (ie, appreciation).

All statistical analyses were conducted using SPSS 20.0, applying a significance level of .05 for single variables and .10 for interaction terms [53].

Results

Study Sample, Baseline Differences, and Attrition Analysis

The CONSORT-EHEALTH flowchart [34] (Figure 3) shows the number of participants who were randomly assigned to one of the 3 study conditions as well as their flow through the study. In total, 1419 participants completed the baseline questionnaire. At 6 months follow-up, data from 1015 (71.53%) participants were collected. In the video condition, only 328 (70.54%) participants had completed the first session of the intervention, whereas 364 (74.13%) did in the text condition. Overall, the average number of completed sessions was 2.15 (SD 1.94) sessions. In total, 10.88% (104/956) of the participants had completed the intervention fully (ie, use of all 6 sessions).

Table 1 provides a comprehensive overview of all baseline characteristics of the study sample, including baseline differences between the 3 study conditions. Participants' mean age was 48.13 (SE 0.31) and 58.56% (831/1419) were female. The mean BMI was 26.42 (SE 0.06), and 73.50% (1043/1419) of the study sample was overweight. The majority had a high level of education (769/1419=54.19%), while fewest participants had a low level of education (214/1419=15.08%). The distribution of educational level differed significantly between the 3 study conditions (Pearson χ^2_4=10.380, *P*=.004). Compared to the text and video conditions, significantly more participants in the control condition had a low educational level. Moreover, the number of participants with a medium level of education was significantly higher in the text and control conditions in comparison to the video condition. Last, participants' mean score on goal setting in the video condition was significantly higher in comparison to the text and control conditions ($F_{2,2415}$=4.740, *P*=.009). No other baseline differences were observed.

Attrition analysis identified several significant predictors of dropout. Participants in the video (OR 2.11, 95% CI 1.48-3.00, *P*<.001) and text conditions (OR 3.23, 95% CI 2.29-4.54, *P*<.001) were significantly more likely to drop out compared to participants in the control condition. Attrition was further significantly higher among participants with a low (OR 2.15, 95% CI 1.46-3.16, *P*<.001) and medium (OR 1.37, 95% CI 1.02-1.85, *P*=.037) educational level in comparison to highly educated participants. Older participants were more likely to complete the follow-up questionnaire (OR 0.97, 95% CI 0.96-0.98, *P*<.001). Finally, participants who had lower levels of self-efficacy to improve their diet (OR 1.36, 95% CI 1.06-1.76, *P*=.016), intention to improve their diet (OR 1.22, 95% CI 1.02-1.46, *P*=.031), and coping planning regarding physical activity (OR 1.31, 95% CI 1.04–1.64, *P*=.022) were significantly more likely to drop out.

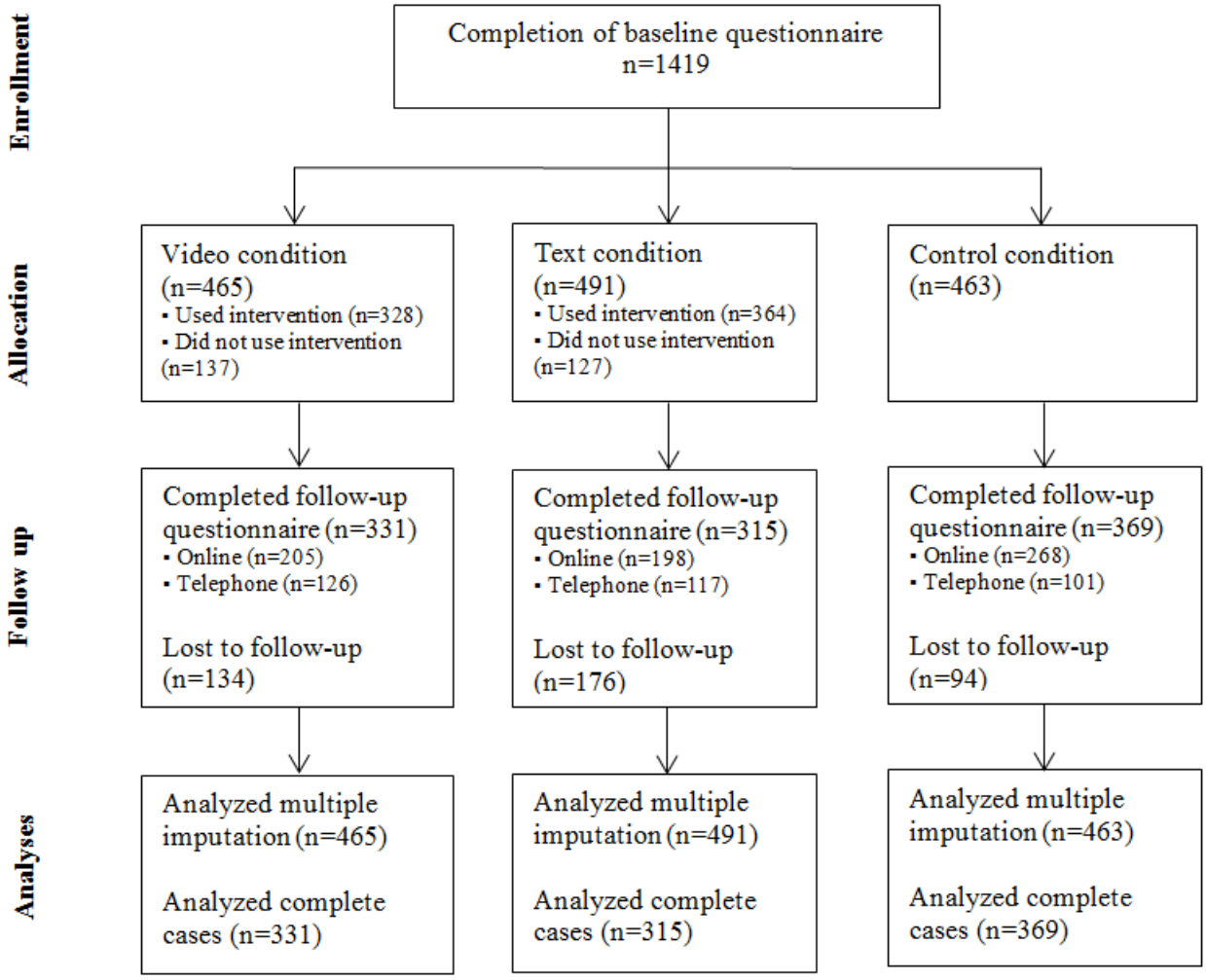
Table 1. Characteristics of the study sample and differences between the study conditions.

	Overall sample (n=1419)	Video (n=465)	Text (n=491)	Control (n=463)	*F* / Pearson χ^2	*df*	*P*
Baseline characteristics							
Gender (female), n (%)	831 (58.56)	273 (58.71)	284 (57.84)	274 (59.18)	0.182	2	.913
Educational level, n (%)					10.380	4	.004[b]
Low	214 (15.08)	75 (16.13)	67 (13.65)[a]	72 (15.55)[a]			
Medium	436 (30.73)	118 (25.38)[a]	161 (32.79)[a]	157 (33.91)[a]			
High	769 (54.19)	272 (58.49)[a]	263 (53.56)	234 (50.54)[a]			
Age, mean (SE)	48.13 (0.31)	48.06 (0.09)	47.84 (0.08)	48.51 (0.08)	0.405	22,415	.667
Self-efficacy improve physical activity, mean (SE)	3.33 (0.02)	3.35 (0.01)	3.35 (0.01)	3.30 (0.01)	0.560	22,415	.571
Self-efficacy improve diet, mean (SE)	3.25 (0.02)	3.28 (0.00)	3.25 (0.00)	3.23 (0.00)	0.831	22,415	.436
Intention improve physical activity, mean (SE)	3.97 (0.03)	3.99 (0.01)	3.97 (0.01)	3.96 (0.01)	0.048	22,415	.953
Intention improve diet, mean (SE)	4.09 (0.03)	4.09 (0.01)	4.12 (0.01)	4.04 (0.01)	0.654	22,415	.520
Goal setting, mean (SE)	3.50 (0.02)	3.60 (0.01)[a]	3.47 (0.01)[a]	3.45 (0.01)[a]	4.740	22,415	.009[b]
Action planning improve physical activity, mean (SE)	3.35 (0.02)	3.34 (0.01)	3.33 (0.01)	3.37 (0.01)	0.214	22,415	.808
Action planning improve diet, mean (SE)	3.22 (0.02)	3.24 (0.01)	3.19 (0.01)	3.25 (0.01)	0.690	22,415	.502
Monitoring, mean (SE)	3.32 (0.02)	3.31 (0.01)	3.30 (0.01)	3.36 (0.01)	0.655	22,415	.520
Coping planning improve physical activity, mean (SE)	3.37 (0.02)	3.35 (0.00)	3.36 (0.01)	3.40 (0.01)	0.745	22,415	.475
Coping planning improve diet, mean (SE)	3.33 (0.02)	3.32 (0.01)	3.32 (0.01)	3.34 (0.01)	0.061	22,415	.941
BMI, mean (SE)	26.42 (0.06)	26.43 (0.02)	26.45 (0.02)	26.37 (0.02)	0.131	22,348	.878
Average daily energy-intake, mean (SE)	1296.91 (13.40)	1308.36 (3.56)	1314.70 (3.51)	1266.51 (3.75)	1.325	22,420	.266
Average daily minutes moderate and vigorous physical activity, mean (SE)	78.23 (2.21)	74.43 (0.53)	76.84 (0.57)	83.52 (0.69)	1.481	22,378	.228
Follow-up characteristics							
BMI, mean (SE)	26.07 (0.08)	25.94 (0.02)	26.11 (0.02)	26.15 (0.02)			
Average daily energy-intake, mean (SE)	1072.57 (20.79)	1016.45 (3.56)	1032.77 (3.60)	1170.70 (3.56)			
Average daily minutes moderate and vigorous physical activity, mean (SE)	107.73 (5.71)	103.17 (0.80)	108.39 (0.80)	111.77 (0.88)			

[a]Values within a row with identical letters were significantly different as determined by analyses of variance with Tukey post-hoc test (for continuous variables) or chi-square tests with Bonferroni correction (for categorical variables).

[b]$P<.05$.

Figure 3. Flowchart of the enrollment, allocation, and participation of respondents.



Intervention Effects on Body Mass Index, Dietary Intake, and Physical Activity

There were no significant interaction effects between type of study condition and educational level for any of the outcome measures.

The regression analyses without interaction terms showed several main intervention effects (Table 2). The video intervention had resulted in a significantly lower BMI compared to the control condition (B=-0.25, *P*=.049), with a small Cohen's *d* effect size of 0.10 [52]. No significant difference was found between the text and control condition regarding BMI (B=-0.09, *P*=.474). Moreover, both the video (B=-175.58, *P*<.001) and text interventions (B=-163.05, *P*=.001) resulted in a significantly lower average daily intake of calories from energy-dense food products compared to the control condition, with medium Cohen's *d* effect sizes of respectively 0.40 and 0.36 [52]. For physical activity, no intervention effects were found for both the video (B=-1.45, *P*=.900) and text conditions (B=1.88, *P*=.863) in comparison to the control condition. In the additional analyses comparing the 2 intervention conditions, no significant differences were found for any of the outcome measures. The complete cases analyses resulted in the same significant findings as the effect analyses with the multiple imputation data described earlier.

Table 2. Intervention effects on the outcome variables at follow-up as assessed by linear regression analyses.

	Video (1) versus control (0) (n=928)[a]					Text (1) versus control (0) (n=954)[a]				
Outcome variables	B[b]	SE	*P*	95% CI	*d*	B[b]	SE	*P*	95% CI	*d*
BMI	-0.25	0.13	.049[c]	-0.50 to 0.00	0.10	-0.09	0.13	.474	-0.35 to 0.16	0.03
Average daily energy-intake	-175.58	45.13	.000[c]	-265.24 to -85.92	0.40	-163.05	48.57	.001[c]	-259.78 to -66.32	0.36
Average daily minutes moderate and vigorous physical activity	-1.45	11.48	.900	-24.28 to 21.38	0.01	1.88	10.88	.863	-19.75 to 23.50	0.02

[a]In the linear regression analyses, the following covariates were included: baseline behavior, educational level, age, goal setting, self-efficacy to improve diet, coping planning regarding physical activity, and intention to improve diet.

[b]B=unstandardized regression coefficient.

[c]*P*<.05.

Process Evaluation

Table 3 provides an overview of the results from the process evaluation. In total, 355 participants completed the process evaluation questionnaire. Overall, the mean scores of the process evaluation variables represented neutral to slightly positive scores for both versions of the intervention, without remarkable low scores. The intervention scored best on usefulness, understandability, and autonomy. The mean score for assessment of the intervention as a whole was 6.85 (SD 1.14).

Regression analyses showed that there was no significant interaction effect of educational level regarding the process evaluation variables. Independent sample *t* tests showed that the information in the video condition was rated as more useful compared to the information provided in the text condition (t_{354}=1.992, *P*=.047). Feelings of relatedness were also significantly higher among participants in the video condition (t_{354}=2.056, *P*=.041) as compared to the text condition. Finally, participants in the video condition rated the intervention significantly better than participants in the text condition (t_{354}=2.388, *P*=.018).

Table 3. Mean and standard deviation of process evaluation variables at follow-up, including differences between the video and text conditions.

Process evaluation variables	Complete cases (n=355)	Video (n=177)	Text (n=178)	*T* (df=354)	*P*
The feedback messages fit to my own situation	3.36 (0.94)	3.41 (0.94)	3.32 (0.94)	0.865	.387
The feedback messages were understandable	3.88 (0.82)	3.91 (0.82)	3.84 (0.82)	0.771	.441
The feedback messages were useful	3.54 (0.92)	3.63 (0.91)[a]	3.44 (0.93)[a]	1.992	.047[b]
The feedback messages were interesting	3.38 (1.00)	3.47 (1.01)	3.30 (0.98)	1.616	.107
Feelings of autonomy	3.98 (0.75)	4.05 (0.74)	3.90 (0.75)	1.815	.070
Feelings of relatedness	3.04 (1.02)	3.15 (1.01)[a]	2.93 (1.02)[a]	2.056	.041[b]
Feelings of competence	3.15 (0.99)	3.23 (0.95)	3.08 (1.03)	1.497	.135
Overall grade intervention (1-10)	6.85 (1.14)	7.00 (1.15)[a]	6.70 (1.12)[a]	2.388	.018[b]

[a]Values within a row with identical letters were significantly different as determined by independent samples *t* tests.

[b]*P*<.05.

Discussion

Principal Findings

The aim of this study was to examine the effects and appreciation of video and text versions of a Web-based computer-tailored obesity prevention intervention among Dutch adults with low and high levels of education.

Our results showed no significant group × education interaction effects. This implies that both versions of the intervention were equally effective for all educational levels. The video version was the most effective intervention because it resulted in both a lower BMI and lower energy intake (compared to the control condition), while the text version had only a lower energy intake. No intervention effects on physical activity were found. Appreciation of the 2 intervention versions also did not differ per educational level. Yet the video version was appreciated more than the text version on usefulness of messages, feelings of relatedness, and grade given to intervention. Overall, it can be concluded that the video intervention performed better than the text intervention regardless of participants' educational level.

The fact that we did not find support for our hypothesis that the video version would be more effective for people with a low educational level is not surprising. A recent similar study into

a smoking cessation intervention has, for example, also only found a main effect of the video condition and no differential effects per educational level [45]. Furthermore, our hypothesis that video messages may work better for lower educated people [26] was based on indications and assumptions derived from a few previous studies but the evidence for this hypothesis was not compelling. Nevertheless, our study provides preliminary evidence that the use of videos in a Web-based computer-tailored intervention can be effective in the prevention of obesity regardless of people's educational level [45].

Both the video and text versions had the strongest effects on dietary intake, which is a finding in line with 2 reviews on Web-based computer-tailored interventions [12,16]. The medium effect size indicates that this effect is of clinical relevance [52] and may suggest important public health potential when the intervention is implemented at a large scale. In line with other studies [11,15], only a small effect size was found regarding BMI (of the video version). However, even small intervention effects on BMI can have a large public health impact resulting in a significant reduction of many health problems, an improved quality of life, and cost savings [54-57]. The fact that the text version did not have a significant effect on BMI can possibly be explained by the fact that the effect size for dietary intake (0.36) was somewhat smaller compared to the effect size of the video version (0.40). The fact that no intervention effect was found on physical activity is not surprising. Many reviews have reported mixed findings of Web-based computer-tailored interventions for physical activity [12,19,20]. One explanation for this finding could be that we encountered problems with the measurement of physical activity in our study. The average daily minutes of moderate-to-vigorous intensity physical activity scores were unrealistically high. Consequently, many participants did not receive the advice to increase their physical activity level within the intervention, resulting in little improvement for this behavior. Recently, this problem has also been identified in a similar efficacy study. Hence, future studies should take this problem into account when assessing physical activity and developing algorithms to deliver tailored messages [46].

In line with two recent studies [21,45], it can further be concluded that a Web-based computer-tailored intervention consisting of videos is appreciated better than an identical intervention that consists of merely text. This difference in appreciation can possibly also explain why the video version was more effective than the text version. The Elaboration Likelihood Model [58], for example, suggests that when information is perceived as interesting and attractive, it is more likely that central route processing will occur. Information that is processed via this central path will have a more long-lasting persuasive effect on the receiver [32]. The better appreciation of the video version may therefore have resulted in more central route information processing. The fact that there were no differential effects in appreciation per educational level can also possibly explain the absence of educational differences in effects. This explanation is supported by the Communication Persuasion Matrix [59], which assumes that effective persuasion is the result of, among others, a suitable media channel. Yet our results demonstrate that the video version (ie, delivery format) is not more attractive for people with a low educational level (ie, user), and therefore no differences in outputs (ie, intervention outcomes) per educational level can be expected.

Strengths and Limitations

An important strength is that this is one of the first studies that has examined whether the use of videos can improve the effectiveness and attractiveness of Web-based computer-tailored interventions. Another strength is that our intervention met several criteria related to higher effectiveness of weight management interventions, such as the use of self-regulation theories [25,37,38], the small changes approach [54], and the Intervention Mapping protocol [36]. A final strength is that we found exactly the same results with the multiple imputation data and the complete cases data.

A limitation of this study is that all outcome measures were self-reported [60,61]. For example, it would have been better to measure BMI objectively. However, research has shown that self-reported BMI does not affect results when used as a continuous variable in a longitudinal study [62]. In addition, the SQUASH resulted in unrealistically high physical activity scores. As these scores were used to provide tailored feedback, it is likely that participants may have received inadequate feedback about their physical activity level. Future interventions should therefore aim to correct for this overestimation in tailoring algorithms or use objective measurements. Moreover, it should be noted that only a relatively small number of participants had completed the process evaluation questionnaire. Yet it is unclear if the inclusion of dropouts would have led to lower or higher scores. For example, participants may have dropped out because they had achieved their goal or because they did not enjoy the intervention. Finally, the relatively short follow-up period of 6 months can be regarded as a limitation. Research with a longer follow-up period is necessary to examine whether or not the effects will be maintained long term.

Conclusions

The video version of the intervention was more effective and better appreciated than the text version, regardless of participants' educational level. Hence, our study provides evidence that the effectiveness of future Web-based computer-tailored obesity prevention interventions can possibly be improved by including videos as a delivery format in tailored health information. Our study shows that this is feasible and effective for Dutch adults with a healthy weight and limited overweight. However, more research is needed to study the long-term effects of the video version.

Acknowledgments

The study was funded by ZonMw, the Netherlands Organization for Health Research and Development (Grant No 200110001).

Conflicts of Interest

HdV is the scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [34].

[PDF File (Adobe PDF File), 673KB - jmir_v17i10e236_app1.pdf]

References

1. Stevens GA, Singh GM, Lu Y, Danaei G, Lin JK, Finucane MM, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index). National, regional, and global trends in adult overweight and obesity prevalences. Popul Health Metr 2012;10(1):22 [FREE Full text] [doi: 10.1186/1478-7954-10-22] [Medline: 23167948]
2. World Health Orgnaization. Obesity: Situation and trends. 2014. URL: http://www.who.int/gho/ncd/risk_factors/obesity_text/en/ [accessed 2014-11-17] [WebCite Cache ID 6U9NzA8LB]
3. World Health Organization. Obesity and overweight. 2014. URL: http://www.who.int/mediacentre/factsheets/fs311/en/ [WebCite Cache ID 6U9OC5yzR]
4. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index). National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet 2011 Feb 12;377(9765):557-567 [FREE Full text] [doi: 10.1016/S0140-6736(10)62037-5] [Medline: 21295846]
5. Bemelmans W, Hoogenveen R, Visscher T, Verschuren W. Toekomstige ontwikkelingen in overgewicht - Inschatting effecten op de volksgezondheid. Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu (RIVM); 2004.
6. van Bakel AM, Zantinge EM. Neemt het aantal mensen met obesitas toe of af?. Bilthoven: RIVM; 2014.
7. McLaren L. Socioeconomic status and obesity. Epidemiol Rev 2007;29:29-48. [doi: 10.1093/epirev/mxm001] [Medline: 17478442]
8. Devaux M, Sassi F. Social inequalities in obesity and overweight in 11 OECD countries. Eur J Public Health 2013 Jun;23(3):464-469 [FREE Full text] [doi: 10.1093/eurpub/ckr058] [Medline: 21646363]
9. Roskam AR, Kunst AE, Van OH, Demarest S, Klumbiene J, Regidor E, et al. Comparative appraisal of educational inequalities in overweight and obesity among adults in 19 European countries. Int J Epidemiol 2010 Apr;39(2):392-404 [FREE Full text] [doi: 10.1093/ije/dyp329] [Medline: 19926664]
10. Uiters E, Verweij A. Overgewicht: Zijn er verschillen naar sociaaleconomische status?. Bilthoven: RIVM; 2012.
11. Wieland LS, Falzon L, Sciamanna CN, Trudeau KJ, Brodney S, Schwartz JE, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. Cochrane Database Syst Rev 2012;8:CD007675 [FREE Full text] [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]
12. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
13. Kayman S, Bruvold W, Stern JS. Maintenance and relapse after weight loss in women: behavioral aspects. Am J Clin Nutr 1990 Nov;52(5):800-807 [FREE Full text] [Medline: 2239754]
14. Klem ML, Wing RR, McGuire MT, Seagle HM, Hill JO. A descriptive study of individuals successful at long-term maintenance of substantial weight loss. Am J Clin Nutr 1997 Aug;66(2):239-246 [FREE Full text] [Medline: 9250100]
15. Broekhuizen K, Kroeze W, van Poppel MNM, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
16. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. Health Educ Res 2009 Aug;24(4):699-720 [FREE Full text] [doi: 10.1093/her/cyp006] [Medline: 19286893]
17. Schulz DN, Smit ES, Stanczyk NE, Kremers SPJ, 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]
18. 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]
19. Enwald HPK, Huotari MA. Preventing the obesity epidemic by second generation tailored health communication: an interdisciplinary review. J Med Internet Res 2010;12(2):e24 [FREE Full text] [doi: 10.2196/jmir.1409] [Medline: 20584698]
20. Neville LM, 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]

21. Soetens KC, Vandelanotte C, de VH, Mummery KW. Using online computer tailoring to promote physical activity: a randomized trial of text, video, and combined intervention delivery modes. J Health Commun 2014 Dec;19(12):1377-1392. [doi: 10.1080/10810730.2014.894597] [Medline: 24749983]
22. 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]
23. Evers KE, Cummins CO, Prochaska JO, Prochaska JM. Online health behavior and disease management programs: are we ready for them? Are they ready for us? J Med Internet Res 2005 Jul 1;7(3):e27 [FREE Full text] [doi: 10.2196/jmir.7.3.e27] [Medline: 15998618]
24. 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]
25. Magnée T, Burdorf A, Brug J, Kremers SP, Oenema A, van AP, et al. Equity-specific effects of 26 Dutch obesity-related lifestyle interventions. Am J Prev Med 2013 Jun;44(6):e57-e66. [doi: 10.1016/j.amepre.2012.11.041] [Medline: 23683991]
26. Walthouwer MJL, Oenema A, Soetens K, Lechner L, De VH. Systematic development of a text-driven and a video-driven web-based computer-tailored obesity prevention intervention. BMC Public Health 2013;13:978 [FREE Full text] [doi: 10.1186/1471-2458-13-978] [Medline: 24138937]
27. van GL, van EP, Boon B, Borsboom G, Visscher T, Oenema A. Results from an online computer-tailored weight management intervention for overweight adults: randomized controlled trial. J Med Internet Res 2012;14(2):e44 [FREE Full text] [doi: 10.2196/jmir.1901] [Medline: 22417813]
28. Brouwer W, Kroeze W, Crutzen R, de NJ, 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]
29. Wilson EAH, Park DC, Curtis LM, Cameron KA, Clayman ML, Makoul G, et al. Media and memory: the efficacy of video and print materials for promoting patient education about asthma. Patient Educ Couns 2010 Sep;80(3):393-398. [doi: 10.1016/j.pec.2010.07.011] [Medline: 20688454]
30. Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR. The prevalence of limited health literacy. J Gen Intern Med 2005 Feb;20(2):175-184 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.40245.x] [Medline: 15836552]
31. Stanczyk NE, Bolman C, Muris JWM, de Vries H. Study protocol of a Dutch smoking cessation e-health program. BMC Public Health 2011;11:847 [FREE Full text] [doi: 10.1186/1471-2458-11-847] [Medline: 22059446]
32. Petty R, Cacioppo J. Communication and persuasion: Central and peripheral routes to attitude change. New York, NY: Springer-Verlag; 1986.
33. Sweller J. Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction 1994 Jan;4(4):295-312. [doi: 10.1016/0959-4752(94)90003-5]
34. 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]
35. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. Int J Behav Nutr Phys Act 2009;6:26 [FREE Full text] [doi: 10.1186/1479-5868-6-26] [Medline: 19457246]
36. Bartholomew L, Parcel G, Kok G, Gottlieb N, Fernandez M. Planning health promotion programs: An intervention mapping approach. 3rd ed. San Francisco, CA: Jossey-Bass; 2011.
37. Zimmerman B, Boekaerts M, Pintrich P, Zeidner M. Attaining self-regulation: A social cognitive perspective. In: Handbook of self-regulation. San Diego, CA: Raven Press; 2000.
38. Leventhal H, Leventhal EA, Contrada RJ. Self-regulation, health, and behavior: A perceptual-cognitive approach. Psychology & Health 1998 Jul;13(4):717-733. [doi: 10.1080/08870449808407425]
39. de VH, 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]
40. van AP, Brug J, Ronda G, Steenhuis I. The relative validity of a short Dutch questionnaire as a means to categorize adults and adolescents to total and saturated fat intake. J Hum Nutr Diet 2001 Oct;14(5):377-390. [Medline: 11906579]
41. EuroFIR AISBL. Dutch Food Composition Database (NEVO). Brussels: RIVM; 2013. URL: http://www.rivm.nl/en/Topics/D/Dutch_Food_Composition_Database/Access_NEVO_data/NEVO_e_book [accessed 2015-10-15] [WebCite Cache ID 6cJ7W1M2A]
42. Wendel-Vos W, Schuit J. Short Questionnaire to Assess Health Enhancing Physical Activity. In: SQUASH. Bilthoven: Netherlands Institute for Public Health and Environment; 2002.
43. Wendel-Vos GCW, Schuit AJ, Saris WHM, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003 Dec;56(12):1163-1169. [Medline: 14680666]
44. UNESCO Institute for Statistics. International Standard Classification of education 2011. UNESCO Institute for Statistics 2012.

45. 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]
46. Schulz DN, Kremers SPJ, Vandelanotte C, van Adrichem MJG, 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]
47. Carey KB, Neal DJ, Collins SE. A psychometric analysis of the self-regulation questionnaire. Addict Behav 2004 Feb;29(2):253-260. [Medline: 14732414]
48. Brown J, Miller W, Lawendowski L, VandeCreek L, Jackson T. The Self-Regulation Questionnaire. In: Innovations in clinical practice: A source book. Sarasota, FL: Professional Resource Press; 1999:281-289.
49. Deci E, Ryan R. Self-determination. New Jersey: Wiley Online Library; 2010.
50. Bodner TE. What improves with increased missing data imputations? Struct Equ Modeling 2008 Oct 22;15(4):651-675. [doi: 10.1080/10705510802339072]
51. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213. [doi: 10.1007/s11121-007-0070-9] [Medline: 17549635]
52. Cohen J. Statistical power analysis for the behavioral sciences. New York, NY: Routledge Taylor and Francis Group; 1988.
53. Rosnow RL, Rosenthal R. Statistical procedures and the justification of knowledge in psychological science. Am Psychol 1989;44(10):1276-1284. [doi: 10.1037/0003-066X.44.10.1276]
54. Hill JO. Can a small-changes approach help address the obesity epidemic? A report of the Joint Task Force of the American Society for Nutrition, Institute of Food Technologists, and International Food Information Council. Am J Clin Nutr 2009 Feb;89(2):477-484 [FREE Full text] [doi: 10.3945/ajcn.2008.26566] [Medline: 19088151]
55. Finkelstein EA, Khavjou OA, Thompson H, Trogdon JG, Pan L, Sherry B, et al. Obesity and severe obesity forecasts through 2030. Am J Prev Med 2012 Jun;42(6):563-570. [doi: 10.1016/j.amepre.2011.10.026] [Medline: 22608371]
56. Verkleij SPJ, Adriaanse MC, Wendel-Vos GCW, Schuit AJ. Longitudinal relation between weight change and quality of life in a community-based population: a prospective cohort study. Eur J Public Health 2013 Apr;23(2):285-290 [FREE Full text] [doi: 10.1093/eurpub/cks058] [Medline: 22645241]
57. Stevens VJ, Obarzanek E, Cook NR, Lee IM, Appel LJ, Smith WD, Trials for the Hypertension Prevention Research Group. Long-term weight loss and changes in blood pressure: results of the Trials of Hypertension Prevention, phase II. Ann Intern Med 2001 Jan 2;134(1):1-11. [Medline: 11187414]
58. Petty R, Cacioppo J. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205. [doi: 10.1016/S0065-2601(08)60214-2]
59. McGuire WJ. Theoretical foundations of campaigns. Public communication campaigns 1981:43-66.
60. Schwarz N. Self-reports: How the questions shape the answers. Am Psychol 1999;54(2):93-105. [doi: 10.1037/0003-066X.54.2.93]
61. Matthews C. Use of self-report instruments to assess physical activity. In: Welk GJ, editor. Physical activity assessments for health-related research. Champaign, IL: Human Kinetics; 2002:107-121.
62. Dahl AK, Hassing LB, Fransson EI, Pedersen NL. Agreement between self-reported and measured height, weight and body mass index in old age—A longitudinal study with 20 years of follow-up. Age and Ageing 2010;39(4):445-451. [doi: 10.1093/ageing/afq038]

Abbreviations

BMI: Body Mass Index
SQUASH: Short Questionnaire to Assess Health-Enhancing Physical Activity

Edited by G Eysenbach; submitted 27.11.14; peer-reviewed by D Thompson, G O'Reilly, M Hingle; comments to author 25.05.15; revised version received 29.07.15; accepted 18.08.15; published 19.10.15.

Please cite as:
Walthouwer MJL, Oenema A, Lechner L, de Vries H
Comparing a Video and Text Version of a Web-Based Computer-Tailored Intervention for Obesity Prevention: A Randomized Controlled Trial
J Med Internet Res 2015;17(10):e236
URL: http://www.jmir.org/2015/10/e236/
doi:10.2196/jmir.4083
PMID:26481772

Original Paper

An eHealth Application in Head and Neck Cancer Survivorship Care: Health Care Professionals' Perspectives

Sanne Duman-Lubberding[1], MSc; Cornelia F van Uden-Kraan[1,2], PhD; Niels Peek[3], PhD; Pim Cuijpers[2], PhD; C René Leemans[1], MD, PhD; Irma M Verdonck-de Leeuw[1,2], PhD

[1]Department of Otolaryngology / Head and Neck Surgery, VU University Medical Center, Amsterdam, Netherlands

[2]Department of Clinical Psychology, VU University, Amsterdam, Netherlands

[3]Health e-Research Centre, University of Manchester, Manchester, United Kingdom

Corresponding Author:
Sanne Duman-Lubberding, MSc
Department of Otolaryngology / Head and Neck Surgery
VU University Medical Center
PO Box 7057
Amsterdam, 1007 MB
Netherlands
Phone: 31 20 444 0931
Fax: 31 20 444 3688
Email: s.lubberding@vumc.nl

Abstract

Background: Although many cancer survivors could benefit from supportive care, they often do not utilize such services. Previous studies have shown that patient-reported outcomes (PROs) could be a solution to meet cancer survivors' needs, for example through an eHealth application that monitors quality of life and provides personalized advice and supportive care options. In order to develop an effective application that can successfully be implemented in current health care, it is important to include health care professionals in the development process.

Objective: The aim of this study was to investigate health care professionals' perspectives toward follow-up care and an eHealth application, OncoKompas, in follow-up cancer care that monitors quality of life via PROs, followed by automatically generated tailored feedback and personalized advice on supportive care.

Methods: Health care professionals involved in head and neck cancer care (N=11) were interviewed on current follow-up care and the anticipated value of the proposed eHealth application (Step 1). A prototype of the eHealth application, OncoKompas, was developed (Step 2). Cognitive walkthroughs were conducted among health care professionals (N=21) to investigate perceived usability (Step 3). Interviews were recorded, transcribed verbatim, and analyzed by 2 coders.

Results: Health care professionals indicated several barriers in current follow-up care including difficulties in detecting symptoms, patients' perceived need for supportive care, and a lack of time to encourage survivors to obtain supportive care. Health care professionals expected the eHealth application to be of added value. The cognitive walkthroughs demonstrated that health care professionals emphasized the importance of tailoring care. They considered the navigation structure of OncoKompas to be complex. Health care professionals differed in their opinion toward the best strategy to implement the application in clinical practice but indicated that it should be incorporated in the HNC cancer care pathway to ensure all survivors would benefit.

Conclusions: Health care professionals experienced several barriers in directing patients to supportive care. They were positive toward the development and implementation of an eHealth application and expected it could support survivors in obtaining supportive care tailored to their needs. The cognitive walkthroughs revealed several points for optimizing the application prototype and developing an efficient implementation strategy. Including health care professionals in an early phase of a participatory design approach is valuable in developing an eHealth application and an implementation strategy meeting stakeholders' needs.

(J Med Internet Res 2015;17(10):e235) doi:10.2196/jmir.4870

KEYWORDS

cancer; tertiary prevention; participatory design approach; follow-up care; supportive care

Introduction

Many cancer survivors have to manage the adverse effects of cancer and its treatment. Head and neck cancer (HNC) specifically has an impact on survivors compared to other cancers. In addition to symptoms such as fatigue, HNC survivors are confronted with oral dysfunction, voice, speech, and swallowing problems, and related social withdrawal and psychological distress. These may negatively impact on quality of life (QOL) [1,2] and increase the need for supportive care.

Supportive care in cancer entails the prevention and management of the adverse effects of cancer and its treatment across the survivorship continuum [3,4]. Although many cancer survivors, including HNC survivors, could benefit from supportive care, they often do not utilize such services [5-8]. Barriers that stand in the way of obtaining supportive care include a lack of awareness of these services and a lack of identification of survivors' symptoms and supportive care needs [9-11].

The use of patient-reported outcome measures (PROs) has been identified as a possible facilitator to detecting survivors' symptoms [12]. Monitoring symptoms may be helpful in addressing survivors' individual supportive care needs [13]. A prerequisite for its success is that monitoring should be followed by adequate referral to supportive care. An eHealth application integrating PROs to monitor QOL, followed by automatically generated tailored feedback and personalized advice on supportive care options, could be an alternative solution to meet cancer survivors' individual needs. The proposed eHealth application could also be a helpful tool to enhance self-management among HNC survivors.

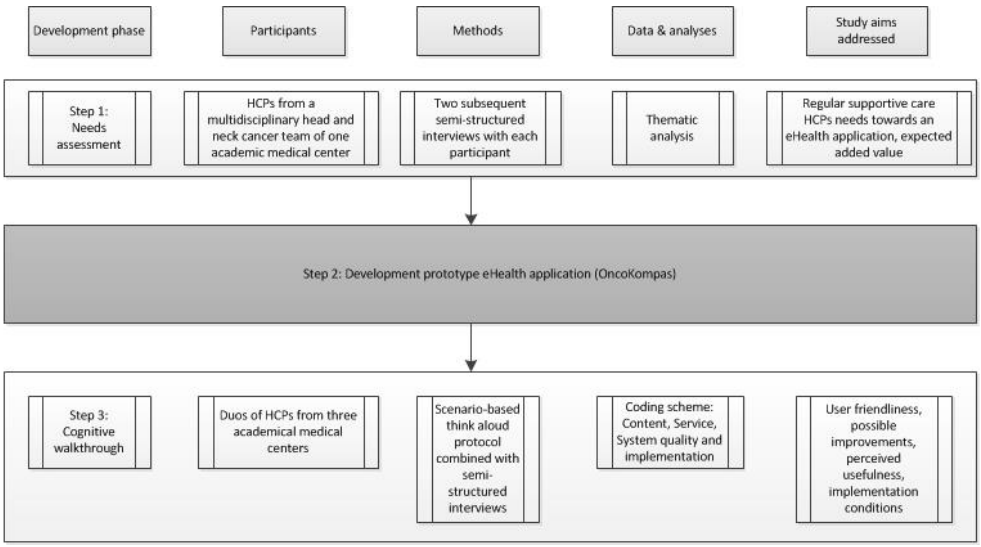
In a previous study, we investigated the attitude and preferences of cancer survivors toward an eHealth application targeting personalized referral to supportive care services [14]. The results of this needs assessment showed that survivors were indeed interested in this option of self-management support and believed that the eHealth application could eliminate barriers experienced in current follow-up care, for example, a minimal response from physicians concerning their needs and having to search for services themselves. The results also highlighted considerations and requirements concerning the application, for example, doubts about the degree of tailoring and the need for the application to be an addition to rather than a substitute for traditional care [14].

In order to develop an effective eHealth application and ensure adequate uptake, it is important to include all stakeholders, including health care professionals, during the entire developmental phase, following an iterative participatory approach [15]. Therefore, the main aim of this study was to investigate health care professionals' perspectives toward an eHealth application in follow-up cancer care, which monitors QOL via PROs (Measure), followed by automatically generated tailored feedback (Learn), and personalized advice on supportive care (Act). The results of this study are intended to contribute to further development of a participatory design approach enabling the development of effective eHealth applications that meet stakeholders' preferences and needs.

Methods

A mixed methods study design was used consisting of 3 steps (Figure 1). We investigated health care professionals' perspectives toward current follow-up care and toward the proposed eHealth application (Step 1) through a qualitative needs assessment. Next, we developed a prototype of the eHealth application (Step 2). Subsequently, we evaluated the application by means of cognitive walkthroughs (CWs) by health care professionals and investigated health care professionals' opinions about usability and conditions for implementation (Step 3).

Figure 1. Study design.



Step 1: Needs Assessment

Health care professionals (N=11) were recruited from a multidisciplinary team involved in the care of HNC patients at the VU University Medical Center in Amsterdam, The Netherlands. We made use of purposive sampling. After permission from the department head, we requested study participation from a heterogeneous sample of health care

professionals. The final sample consisted of an oral and maxillofacial surgeon, head and neck surgeon, oncologist, radiation oncologist, medical social worker, physiotherapist, dental hygienist, dietician, speech therapist, and 2 oncology nurses. Participating health care professionals' experience in working with cancer patients ranged from 2 years and 3 months to 25 years (mean 13.38 years). Health care professionals were interviewed twice. An overview of the topics is shown in Table 1.

Table 1. Topics discussed in the needs assessment interviews.

Topic	Example question
Current follow-up care: assessing symptoms and supportive care needs	How do you assess patients' symptoms and quality of life?
	What difficulties do you encounter when assessing patients' symptoms and quality of life?
	How do you assess patients' supportive care needs?
	Do you refer patients to supportive cancer care options?
	To which supportive care options do you refer patients?
	What difficulties do you encounter when referring patients to supportive cancer care options?
Added value of an eHealth tool in follow-up care for health care professionals	How may an eHealth application be supportive/fit into in your current role in follow-up cancer care?

The first interview covered questions about current follow-up care (assessing patient's symptoms and need for supportive care). The second interview covered questions about the expected added value for health care professionals of an eHealth tool aimed at improving supportive care. In this second interview, more information about the proposed application was conveyed (eg, examples of personalized advice texts and supportive care options).

In total, 22 interviews were conducted, which lasted between 24 and 50 minutes (median 35, SD 7.24). All interviews were recorded and transcribed verbatim.

Step 2: Development of the Prototype eHealth Application

A prototype of the eHealth application, "OncoKompas," was developed based on the results of the needs assessment among health care professionals (from this study) and survivors [14]. Existing applications were used as examples to build the application [16,17]. First, the results of both needs assessments were discussed with the development team (Web designers and programmers), to translate these needs into requirements. The Web designer and programmers used their expertise to translate these requirements into a prototype of OncoKompas. During regular "demo sessions," these requirements were revisited to ensure a proper translation into the prototype. The contents of OncoKompas were developed together with teams of experts consisting of cancer survivors, medical specialists, and paramedics (refer to the "Results" section for more details on OncoKompas).

Step 3: Cognitive Walkthroughs

The cognitive walkthroughs (CWs) consisted of an expert-based usability evaluation followed by semistructured interviews. The health care professionals who participated in the needs assessment in Step 1 were complemented by a psychologist, a spiritual counsellor, and a patient advisor. We also included 3 head and neck surgeons, a radiation oncologist, 2 oncology nurses, and a health scientist from 2 other academic hospitals.

All but one of the usability evaluations were conducted in pairs of health care professionals because this was expected to increase "thinking out loud" by the participants. Health care professionals were asked to "walk through" the application guided by scenarios and user tasks from the end-users' viewpoint. Following the usability evaluations, we interviewed the health care professionals on the implementation process (Table 2).

Table 2. Overview of CW scenario's tasks and interview topics.

Scenario example	This scenario involves a 66-year-old female head and neck cancer patient. She is experiencing (the onset of) depression as well as stress at home. Furthermore, she has diarrhea and does not use a feeding tube or nutritional drinks. She has mild dysphagia and moderate loss of taste and smell.
CW tasks	Task 1: Monitor disease problems by filling out the PROs in OncoKompas and sending in the completed questionnaires.
	Task 2: View your personal well-being profile in OncoKompas.
	Task 3: Use personalized well-being profiles to find information regarding your physical condition related to your tumor.
	Task 4: Find personalized advice on an aspect of interest to you, and then take action based on this advice.
	Task 5: Find more information in OncoKompas regarding a particular supportive care option of your choice and then open and view the website of a recommended supportive care provider.
Semistructured interview topic: Implementation OncoKompas	What role do you think you could have in the usage of OncoKompas by patients?
	How do you think OncoKompas could be implemented in the regular follow-up care procedure?
	Do you intend to refer your patients to OncoKompas when available?

In total, 11 CWs were conducted, which lasted between 68 and 120 minutes (median 82, SD 14.54), and were recorded using Morae software (Morae version 2.1, TechSmith).

Data Analysis

All needs assessment interviews and CWs were analyzed by thematic analysis [18]. Both coders (SDL and CvU) read all transcripts to familiarize themselves with the data. The coders independently selected citations from the transcripts of all needs assessment interviews relating to current follow-up cancer care and needs of health care professionals with respect to an eHealth application. These were coded into themes.

To analyze the usability of OncoKompas, we made use of the CW transcripts, supported by the Morae recordings. In total, 9 transcripts were coded by 2 coders. Initial codes for the CWs were generated focusing on system quality (ease of use), content quality (usefulness and relevance), and service quality (the process of care provided) [15,19,20]. Additionally, both coders independently selected citations for 9 of the semistructured interviews concerning the implementation process and coded these into categories. The remaining 2 CW transcripts were coded by 1 coder (SDL).

Next, the 2 coders met to review the extracted citations and themes from the needs assessment interviews and CWs. Disagreements were resolved through consensus, which was reached on all citations and themes. They developed 2 frameworks (one for the needs assessment and one for the CWs), in which the themes were identified and subthemes defined. After coding, the raw data were examined again to ensure the robustness of the analytical process and to ensure that all the data were reflected in the coding [21]. Quotations were translated from Dutch into English and anonymized.

Results

Step 1: Health Care Professionals' Needs Assessment

Current Follow-Up Care: Assessing Symptoms and Supportive Care Needs

Health care professionals indicated that during consultations they typically ask the cancer survivors about their symptoms and undertake a physical examination. A few indicated they also asked their patients to complete PROs. Furthermore, when preparing for the consultations, health care professionals indicated that they consulted with their colleagues, as well as the electronic hospital information system (Table 3).

Health care professionals mentioned several difficulties in assessing survivors' symptoms and in the referral process to supportive care services. They mentioned they are able to address only a limited scope of issues during the consultation due to limited time. In addition, all health care professionals said they tend to focus on their own field of expertise, for example, physicians indicated that they do not feel capable of assessing a survivor's psychological well-being:

> *In an open setting you will of course ask: "Are there any things you'd like to discuss?" I think that works fine as a first move to also allow space for the psychological aspects, but of course you do ask things like: "How is your weight?," "What about the pain?."*

Another difficulty according to health care professionals is that they do not want to burden the survivor with unnecessary questions about irrelevant or irreversible symptoms, for example, problems with salivary glands due to radiation therapy. In addition to this, they indicated that they lack a complete picture of survivors' symptoms and quality of life. This comes about due to survivors' hesitancy in mentioning all their symptoms and issues, as well as due to fragmentation in clinical care (eg, no insight into the patient information system of the other health care professionals involved).

Table 3. Overview of key issues and themes from the needs assessment.

Key issues	Themes
Detecting symptoms and need for supportive care	
Assessment of survivors' symptoms	
Consulting survivor	Verbal questioning (based on checklist or according to protocol)
	Observing and physical examination (according to protocol)
	Wait and see what symptoms survivor describes
	Use of PROs (OncoQuest)
Consulting colleagues	
Consulting patient information system	
Barriers in determination	
Limited scope of issues being raised during consultation	Limited consultation time
	Limited skills or expertise of health care professional
	Limited responsibility of health care professional
Do not wish to burden the survivor by asking about…	Irrelevant symptoms
	"Irreversible" symptoms caused by treatment
No complete picture of a survivors' symptoms	Patients do not mention all symptoms
	Fragmentation in care
Current referral to supportive care options	
Supportive care services referred	
Available services within the hospital	Allied health services, ie, physical therapist, dietician
Services outside hospital	Specialized cancer centers
	Cancer rehabilitation program
	Allied health services in the region
	General practitioner
Barriers in referral	
Lack of options	Lack of overview of available and adequate supportive care
Practical barriers in referral	Lack of time to encourage survivors to obtain supportive care
	Referral to region complicated due to lack of expertise on HNC
	Referral only possible through physician
No need of survivor to be referred	Survivor is unwilling to be referred
	Survivor already has adequate supportive care
Expected added value of eHealth application in follow-up care for health care professionals	
Increases insight into symptoms	

Key issues	Themes
Provides a complete picture of patients' symptoms	Provides insight into the interdependence of patients' symptoms
	Signal function: creates awareness of the severity of symptoms
	In support of their own observation/impression of health care professional
	By monitoring symptoms ability to serve as treatment outcome
Improved (preparation for) consultation	Low threshold to speak up about specific issues/symptoms
	Option to target questions regarding specific symptoms
	Option to elaborate on and prioritize symptoms
Personalized advice/information	
Provides tailored information	More detailed information than provided by physician
	Back up for advice provided by health care professional
	Supportive to information provided by health care professional
Platform to deliver additional care	Informative support to self-management advice
	Availability of physical therapy exercises
Increases insight into QOL domains	Improved knowledge in QOL domains out of health care professionals' expertise
Insight into supportive care options	Increased insight into supportive care options
Additional service in follow-up care	Showcase for hospital

Health care professionals indicated that care for HNC patients in The Netherlands is provided by multidisciplinary teams during treatment. However, follow-up care is generally provided only by physicians who continue to follow-up on the cancer survivor regularly. Physicians said they were hesitant to refer survivors to supportive care. In cases of mild symptoms, they provide the survivor with personal advice themselves. Where there are cases of severe symptoms, they refer survivors to other health care professionals. The supportive care services that health care professionals make their referrals to are often limited to other health care professionals in the same hospital. When referral takes place to services outside the hospital, these mainly include specialized centers for cancer survivors, cancer rehabilitation programs, allied health services in the region, or the survivor's general practitioner.

Health care professionals also described barriers in referral to supportive care. They reported a lack of overview of the availability of supportive care services. Also, practical barriers in referral were mentioned, including a lack of time to encourage survivors to obtain supportive care:

> *What I usually do, is just say "this is available," and if it will do some people good, they will give it a go if I want them to. In itself, that's fine, but it is tricky, as you only have a short amount of time during a consultation. You have to encourage people too and that is often the problem.*

Referral to allied health services in the region was considered complicated due to a lack of expertise in HNC. Finally, health care professionals indicated there was a perceived lack of need by the survivor to be referred, either due to unwillingness or due to the health care professionals' assumption that the survivor already had adequate support.

Health Care Professionals' Views on a Proposed eHealth Application in Follow-Up Cancer Care

Most health care professionals expected an eHealth application to provide added value for themselves in their practice, particularly in terms of follow-up care with the aim of optimizing supportive cancer care (Table 3). They hoped by using an eHealth application such as this to monitor survivors, to obtain an increased insight into these patients' symptoms. In addition, the application could help detect survivors with severe symptoms. Health care professionals indicated that they anticipated the application could serve as a tool during their consultations, help prioritize symptoms, and support them in elaborating on and targeting questions toward symptoms.

Personalized information and advice for survivors provided by an eHealth application was expected to have an added value, if tailored to tumor type or treatment. Health care professionals indicated they expected this information to be supportive or supplementary to the information they provided to patients. Health care professionals expected that the application could also serve as a platform to deliver additional care, such as self-management advice and physical therapy exercises. Another benefit expected was an increased insight into various quality of life domains that were not part of the health care professionals' specialty.

Insight into supportive care options available could be improved by means of an eHealth application. Finally, health care professionals expected the application to be an additional service for survivors in follow-up care, which could serve as a showcase for the hospital.

Step 2: Prototype of OncoKompas

The prototype OncoKompas was developed in Step 2. OncoKompas was developed as an online computer application.

It consists of the following 3 components: (1) Measure, (2) Learn, and (3) Act. In the "Measure" component, cancer survivors can independently complete PROs targeting the following QOL domains: physical functioning, psychological functioning, social functioning, healthy lifestyle, and existential issues. A specific domain containing topics for head and neck cancer patients is available, in addition to those general domains for cancer survivors (Table 4).

Table 4. Overview of OncoKompas topics.

Psychological QOL	Physical QOL	Social QOL	Healthy lifestyle	Existential issues	Head and neck cancer
Anxiety and depression	General everyday life	Social life	Alcohol	Life questions	Swallowing
Fear of recurrence	Pain	Relationship with partner	Physical activity	Religion	Speech
Subjective cognitive functioning	Sexuality	Relationship with children	Dietary intake	Future perspective	Oral function
Stress	Sleep quality	Financial circumstances	Weight		Neck and shoulder function
	Body image	Patient-physician communication	Smoking		Loss of smell and taste
	Fatigue	Return to work			Head and neck cancer specific lymphedema
	Diarrhea				Nutritional drink/Tube feeding
	Lack of appetite				
	Dyspnea				
	Nausea or vomiting				
	Constipation				
	Hearing and tinnitus				

On the basis of the interview results, specific PROs, validated questionnaires (or subscales) if available, were selected by the project team in collaboration with teams of experts. This selection was based on Dutch practice guidelines and literature searches. Data from the "Measure" component are processed in real-time and linked to tailored feedback to the cancer survivor in the "Learn" component. All algorithm calculations are based on available cutoff scores or are defined based on Dutch practice guidelines, literature searches, and/or consensus by teams of experts. A compass metaphor is used in the "Learn" component to summarize overall well-being. Once overall well-being has been presented, feedback is provided to the participant on the risk level for the topics (eg, depression, fatigue) by means of a 3-color system: green (no elevated well-being risks), orange (elevated well-being risks), and red (seriously elevated well-being risks). Cancer survivors receive elaborate personalized information on the outcomes. For instance, taking depression, information is provided on the symptoms of depression and the proportion of cancer survivors who suffer from depressive symptoms. Special attention is paid to evidence-based associations between outcomes. For example, feedback on the association between depression and fatigue is provided if a participant has an orange or a red score on depression as well as on fatigue. The feedback in the "Learn" component concludes with comprehensive self-care advice (tips and tools). All this advice is tailored to the individual cancer survivor, for example, tailored to age (eg, survivors over 70 years of age receive an adapted advice on exercising), gender (eg, advice on sexuality issues differ between men and women), and comorbidity (eg, dietary advice differs for diabetic patients).

In the "Act" component, survivors are provided with personalized supportive care options based on their PRO scores and expressed preferences (eg, preference for individual therapy versus group therapy). If a participant has elevated well-being risks (orange score), the feedback includes suggestions for self-help interventions. If a participant has "seriously elevated well-being risks" (red score), the feedback includes advice to contact their own medical specialist or general practitioner. If survivors want to share their results with their caregiver, they are able to "print their results to PDF" and either bring these with them (hard copy) during their consultation with the caregiver or email these results to the caregiver.

A clickable demo of the application (in Dutch) or an animation video (in Dutch and English) is available on the OncoKompas website.

Step 3: Cognitive Walkthroughs

Technical errors occurred in 2 of the 11 CWs but were subsequently resolved. Health care professionals' strengths and weaknesses concerning quality of the system, content, and service are presented in Multimedia Appendix 1.

System Quality

Health care professionals' opinions toward the accessibility of OncoKompas varied. Many health care professionals indicated that OncoKompas may not be useful for a group of HNC

survivors, due to limited eHealth literacy skills, lack of motivation, and older age. Others emphasized the usefulness of eHealth applications for HNC survivors, through the elimination of social barriers, such as difficulty speaking and shame about facial scarring. The 24/7 availability from home was considered important (see Multimedia Appendix 1).

According to the majority of health care professionals, the ease of use of OncoKompas was suboptimal because of the complicated navigation structure. Health care professionals mentioned that the interface was too busy for the target group. Complicating aspects included too much scrolling and unclear progress in the "Measure" component of OncoKompas. Positive aspects included a self-explanatory walkthrough of the application and the option to quit and save the questionnaire halfway through.

Health care professionals suggested the level of tailoring needed to be improved, for example, with respect to the advice provided. They considered the advice as distant and general, which could make it unclear to the survivor that the information had been tailored to their situation. Some mentioned that as participants are forced to monitor all symptoms, they might receive information on symptoms irrelevant to them. The provision of tailored advice, in contrast to surfing the Web was considered positive:

> *As I see it, the advantage of the program is that it makes the piles of available information accessible.*

Finally, health care professionals suggested including reminders to encourage participants into action.

Content Quality

Health care professionals believed there was tension between the application goal and the use of evidence-based PROs. The use of evidence-based PROs requires participants to fill out more questions than needed to obtain personalized advice. However, the evidence-based content of the application was valued by health care professionals. The application mostly followed health care professionals' own professional standards with respect to enquiring about symptoms and the provision of advice, which dovetailed with their advice to potential end users. In other words, the advice given in the application was the advice that health care professionals expected to be provided (see Multimedia Appendix 1):

> *When I look at it, it provides the advice that I would expect to be provided.*

Health care professionals varied in their opinion regarding content comprehensibility. The "Measure" component was considered difficult by some health care professionals, as was the use of abstract terminology (eg, "well-being profile"). Others were positive about the different comprehensibility levels at which information was provided to participants. They complimented the formulation of advice texts and the different levels of information provided by the application (so-called read more options):

> *I believe that most people are able to gauge their own level pretty well. And people who cannot fully grasp this information, soon think, well, I have read all the tips, that will do.*

The content was considered to be complete by most health care professionals. They were positive about the completeness of the QOL aspects included, their interdependence, and the diversity in supportive care options provided. Others indicated that the content was superfluous in that some information is provided to participants several times throughout OncoKompas. Some believed information was missing, for example, costs of supportive care options.

Service Quality of OncoKompas

Health care professionals were positive about the usefulness of OncoKompas in identifying symptoms, especially by providing patients with a complete picture of their well-being and insight into the interdependences, leading to a clarification of request for help:

> *OncoKompas is useful...by broadening the insight of patients and clarifying to them when the time has come to ask for help. Instead of having just us as health care professionals ask and explore, it can enable patients to become more pro-active in that respect.*

Health care professionals' concerns included that OncoKompas lacks nuance and may not be as tailored as a personal consultation with a health care professional (see Multimedia Appendix 1).

Health care professionals also indicated that they expected benefits in informing participants by creating an opportunity to receive information on sensitive topics. Others mentioned that survivors may be reluctant to use the application for information, because it is easier for them to contact the outpatient clinic. Some health care professionals expected that participants might receive inaccurate or irrelevant information if they inaccurately navigated through the application:

> *Well, with only a few wrong clicks, you can end up with the strangest of information. That does worry me a bit.*

Health care professionals indicated that the application could support participants by referring them to appropriate supportive care options compatible with their symptoms. Some health care professionals mentioned concerns regarding whether participants would know what to do next. They expected participants to get lost in the supportive care options available to choose from, possibly leading to a lack of action.

Health care professionals also mentioned to expect some overall benefits for future participants, such as empowerment and increased engagement:

> *I can imagine this patient is wondering, "Do I have to bother my physician about that?" And when she receives the information from OncoKompas, she sees, "Yes, I should bother my physician about that."*

According to health care professionals, the application could also help participants be better prepared for their consultations. Concerns mentioned by health care professionals included an expected increase in workload and more consultations with

health care professionals due to participants' increased insight into whom to turn to with their symptoms. Some health care professionals mentioned that OncoKompas could possibly lead to participants' continuing to obsess about their disease instead of helping them move on with their life or that emotions surrounding their cancer could (re)surface. Another negative consequence mentioned by health care professionals was that participants might not seek the expertise of a health care professional concerning their symptoms if they had already received information from the application.

Implementation of OncoKompas

Most health care professionals mentioned a positive intention to refer their patients to OncoKompas. All health care professional agreed that if the application were to be implemented in daily clinical practice, it should be offered to survivors through a routine procedure in a care pathway. Physicians believed that referral to the application should take place from different sources, including outside the hospital (eg, by the Dutch Cancer Society). Health care professionals suggested possibilities to increase awareness, such as providing a demo in the waiting room (see Multimedia Appendix 2).

Health care professionals differed in their opinion toward the best strategy to implement the application in clinical practice. Several health care professionals believed that OncoKompas should be implemented as a self-management instrument (independent use by survivors), while others stressed the use as a *supported* self-management instrument (with support from a health care professional).

Implementation as a self-management instrument was expected to stimulate survivor empowerment and to support survivors in defining their own route to relieving their symptoms and increasing their quality of life. Furthermore, health care professionals mentioned that survivors are responsible for their own well-being. Health care professionals indicated that referral of survivors to their physician by means of OncoKompas in case of severe symptoms would relieve them from the responsibility to take action on symptoms they may not know are present. Health care professionals argued that with a self-management application, survivors' privacy would remain intact. They expected survivors to answer more truthfully if they knew their physician would not have access to the data:

> *When a patient wants to share their results, that would be nice, but I think the additional value also lies in that he has the opportunity to keep it to himself.*

Finally, health care professionals expected that they could get around difficulties in discussing OncoKompas results during their regular consultations (eg, difficulties due to time pressure and the priority to check for cancer recurrence) by offering HNC survivors access to OncoKompas as an unsupported self-management application.

Other health care professionals indicated that OncoKompas should be implemented as a *supported* self-management tool because the responsibility of survivors' well-being always remains with the health care professional. Health care professionals wished to receive feedback through access to OncoKompas or a system alert. They wanted to use the results to discuss these and prioritize symptoms during their consultations. Health care professionals indicated they were aware that when OncoKompas is implemented as a supported self-management tool, this requires action from the health care professionals in cases where survivors receive negative results from the application. Health care professionals mentioned that they might not always be able to fulfil this expectation, possibly leading to survivor disappointment:

> *It might be that it raises false expectations in the patient. As surely there will be times that I won't come round to it and if the patient then expects, the doctor will have a quick read when I am there and we are going to discuss what I have filled in, then that is a bit hard on the patient.*

Discussion

Preliminary Findings

This study investigated health care professionals' perspectives toward current follow-up care and the added value of an eHealth application monitoring QOL via patient-reported outcomes (PROs; Measure) followed by automatically generated tailored feedback (Learn), and personalized advice on supportive care (Act).

Barriers in Referral to Supportive Care and Health Care Professionals' Acceptance of an eHealth Application

The results of this study showed that current referral to optimal supportive care is limited due to several barriers, such as limited consultation time and a lack of overview of supportive care options. Our data support previous studies that have obtained insights into these barriers [7-9,22,23]. Furthermore, health care professionals clearly indicated they expected survivors to mention their symptoms. However, previous studies have shown that survivors themselves also experience barriers possibly resulting in unmet needs [8,10,14]: emotional barriers, such as not wanting to complain after surviving cancer, and practical barriers, such as not wanting to burden their physician. By automating the referral process to supportive care by means of an eHealth application, a barrier such as not wanting to burden their physician may be removed. In general, health care professionals expected that the proposed eHealth application could optimize the referral to supportive care.

Content, System, and Service Quality of OncoKompas

Overall, health care professionals were most pleased with the service quality of the application but mentioned several considerations regarding its system and content quality.

Our study showed that health care professionals concluded that OncoKompas was useful for a limited group of (HNC) survivors. A frequently mentioned barrier was lack of Internet access, which is remarkable as a large majority of the Dutch population (90.4%) has access to the Internet; 80% of 65-75 year olds indicated they used the Internet [24]. Therefore, access to the Internet seems to have become less of a barrier and the emphasis should be on developing an application congruent with eHealth literacy skills of end users. The needs assessment among cancer

survivors showed that they required the application to be easily comprehensible [14].

Health care professionals in our study underlined the importance of tailoring the application. In the needs assessment among cancer survivors, tailoring was also deemed important. Patients, however, mentioned doubts about the degree of tailoring that is possible [14]. According to health care professionals, a lack of tailoring could lead to a loss of interest, possibly leading to nonadherence [25]. Health care professionals suggested that only select topics of concern to a user should be provided to improve tailoring.

Considering the content quality of OncoKompas, the majority of issues mentioned were related to the use of PROs. Usage of PROs resulted in overlap between items (as individual items cannot be deleted from validated PROs). Additionally, health care professionals mentioned comprehensibility issues: they assessed several PROs as too difficult. Although we strived for readability at the 10thgrade level in all texts in the "Learn" and "Act" components of OncoKompas, validated PROs are not always at this reading level.

This study demonstrated that most health care professionals expect that the application will support survivors in obtaining appropriate and timely supportive care tailored to their symptoms [14]. This is in line with results of the needs assessment among cancers survivors. They expected similar advantages in receiving information on supportive care options tailored to their specific needs, for example, the ability to find supportive care options on their own and to take actions toward their symptoms [14]. In directing the HNC survivor to optimal supportive care, OncoKompas meets the objective of the current cancer care navigation movement toward ensuring cancer survivors receive adequate follow-up and supportive care [26,27]. However, some health care professionals in our study doubted whether survivors would know what to do after completing OncoKompas. They expected that HNC survivors could get lost in the supportive care options they can choose from, possibly leading to a lack of action. Given the evidence that more options and choice equals more stress and less action [28], the number of supportive care options that OncoKompas offers to the participant is limited to 3 recommendations.

Implementation of eHealth

Health care professionals differed in their opinion whether OncoKompas should be implemented as a self-management application or a *supported* self-management application. The consequences of implementation on existing working procedures were discussed in interviews with those who preferred to implement OncoKompas as a *supported* self-management application, for example, incorporating an alert system in the hospital patient information system. Other health care professionals were of the opinion that survivors are responsible for their own well-being and that because of the importance in empowering the survivor and respect for the survivor's privacy, the application should be implemented as a stand-alone self-management instrument. Wiggers et al [29] reported that implementing a *supported* self-management eHealth application in routine clinical practice increases the complexity of existing working procedures, possibly leading to low uptake of an eHealth application. This barrier may be avoided when implementing the application as a self-management tool. Both options offer advantages in clinical practice: supported self-management applications may be more suitable for survivors who lack eHealth literacy skills, while other cancer survivors may be more empowered by a stand-alone self-management instrument. Consequences of both options need to be studied further.

Strengths and Limitations

This study is limited due to the small number of health care professionals involved. Another limitation is that it might have been difficult for health care professionals to view an eHealth application from the survivors' perspective. However, the use of a participatory design approach, including health care professionals from different academic hospitals as well as combining these results with cancer survivors' perspectives [14], covered all main aspects. The added value of usability research is limited when weaknesses are mentioned that could have been prevented in the design process. Because there are no similar applications in oncology, the results of our study add value and can be used as a guide for designing other applications. A strength of this study is that we also gained insight into implementation requirements of eHealth in clinical practice.

Conclusion

Health care professionals experienced a variety of barriers in the current organization of supportive cancer care, such as a lack of overview of options. Health care professionals expected that the use of an eHealth application that monitors QOL and provides automatically generated personalized advice and referral to supportive care options may be helpful in eliminating some of these barriers. However, they also highlighted some concerns. They mentioned that the application may not be useful for all HNC survivors due to limited eHealth literacy and an older age. Cognitive walkthroughs revealed several points for optimizing the prototype of the application, including improved tailoring. Health care professionals expected several advantages for survivors: insight into the interdependence of symptoms for cancer survivors, (earlier) referral to adequate supportive care, and increased patient empowerment. Finally, useful recommendations for developing an efficient implementation strategy appeared from the interviews. It can be concluded that including health care professionals in an early phase of a participatory design approach is valuable in designing an eHealth application and an implementation strategy that meets stakeholders' needs.

Acknowledgments

We would like to thank the health care professionals who contributed to this study and Nynke Dijkstra and Alexander Scheffer for their contribution to the data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1

System, content, and service quality OncoKompas (strengths and weaknesses).

[PDF File (Adobe PDF File), 206KB - jmir_v17i10e235_app1.pdf]

Multimedia Appendix 2

Implementation of OncoKompas.

[PDF File (Adobe PDF File), 109KB - jmir_v17i10e235_app2.pdf]

References

1. de Graeff A, de Leeuw JR, Ros WJ, Hordijk GJ, Blijham GH, Winnubst JA. Long-term quality of life of patients with head and neck cancer. Laryngoscope 2000 Jan;110(1):98-106. [doi: 10.1097/00005537-200001000-00018] [Medline: 10646723]
2. Verdonck-de Leeuw IM, Eerenstein SE, Van der Linden MH, Kuik DJ, de Bree R, Leemans CR. Distress in spouses and patients after treatment for head and neck cancer. Laryngoscope 2007 Feb;117(2):238-241. [doi: 10.1097/01.mlg.0000250169.10241.58] [Medline: 17277616]
3. Hodgkinson K, Butow P, Hunt GE, Pendlebury S, Hobbs KM, Wain G. Breast cancer survivors' supportive care needs 2-10 years after diagnosis. Support Care Cancer 2007 May;15(5):515-523. [doi: 10.1007/s00520-006-0170-2] [Medline: 17120068]
4. Surbone A, Peccatori F. Unmet needs of cancer survivors: supportive care's new challenge. Support Care Cancer 2006 May;14(5):397-399. [doi: 10.1007/s00520-006-0022-0] [Medline: 16604352]
5. Carlson LE, Angen M, Cullum J, Goodey E, Koopmans J, Lamont L, et al. High levels of untreated distress and fatigue in cancer patients. Br J Cancer 2004 Jun 14;90(12):2297-2304 [FREE Full text] [doi: 10.1038/sj.bjc.6601887] [Medline: 15162149]
6. Hack TF, Carlson L, Butler L, Degner LF, Jakulj F, Pickles T, et al. Facilitating the implementation of empirically valid interventions in psychosocial oncology and supportive care. Support Care Cancer 2011 Aug;19(8):1097-1105. [doi: 10.1007/s00520-011-1159-z] [Medline: 21494781]
7. Steginga SK, Campbell A, Ferguson M, Beeden A, Walls M, Cairns W, et al. Socio-demographic, psychosocial and attitudinal predictors of help seeking after cancer diagnosis. Psychooncology 2008 Oct;17(10):997-1005. [doi: 10.1002/pon.1317] [Medline: 18203243]
8. Verdonck-de Leeuw IM, de Bree R, Keizer AL, Houffelaar T, Cuijpers P, van der Linden MH, et al. Computerized prospective screening for high levels of emotional distress in head and neck cancer patients and referral rate to psychosocial care. Oral Oncol 2009 Oct;45(10):e129-e133. [doi: 10.1016/j.oraloncology.2009.01.012] [Medline: 19362038]
9. Eakin EG, Strycker LA. Awareness and barriers to use of cancer support and information resources by HMO patients with breast, prostate, or colon cancer: patient and provider perspectives. Psychooncology 2001;10(2):103-113. [Medline: 11268137]
10. Steele R, Fitch MI. Why patients with lung cancer do not want help with some needs. Support Care Cancer 2008 Mar;16(3):251-259. [doi: 10.1007/s00520-007-0301-4] [Medline: 17602245]
11. Walsh J, Young JM, Harrison JD, Butow PN, Solomon MJ, Masya L, et al. What is important in cancer care coordination? A qualitative investigation. Eur J Cancer Care (Engl) 2011 Mar;20(2):220-227. [doi: 10.1111/j.1365-2354.2010.01187.x] [Medline: 20477854]
12. Kotronoulas G, Kearney N, Maguire R, Harrow A, Di Dominico D, Croy S, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. J Clin Oncol 2014 May 10;32(14):1480-1501. [doi: 10.1200/JCO.2013.53.5948] [Medline: 24711559]
13. Carlson LE, Waller A, Groff SL, Zhong L, Bultz BD. Online screening for distress, the 6th vital sign, in newly diagnosed oncology outpatients: randomised controlled trial of computerised vs personalised triage. Br J Cancer 2012 Aug 7;107(4):617-625 [FREE Full text] [doi: 10.1038/bjc.2012.309] [Medline: 22828610]
14. Lubberding S, van Uden-Kraan CF, Te Velde EA, Cuijpers WJMJ, Leemans CR, Verdonck-de Leeuw IM. Improving access to supportive cancer care through an eHealth application: a qualitative needs assessment among cancer survivors. J Clin Nurs 2015 May;24(9-10):1367-1379. [doi: 10.1111/jocn.12753] [Medline: 25677218]
15. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

16. de Bree R, Verdonck-de Leeuw IM, Keizer A, Houffelaar A, Leemans CR. Touch screen computer-assisted health-related quality of life and distress data collection in head and neck cancer patients. Clin Otolaryngol 2008 Apr;33(2):138-142. [doi: 10.1111/j.1749-4486.2008.01676.x] [Medline: 18429869]
17. Vosbergen S, Laan EK, Colkesen EB, Niessen MA, Kraaijenhagen RA, Essink-Bot ML, et al. Evaluation of end-user satisfaction among employees participating in a web-based health risk assessment with tailored feedback. J Med Internet Res 2012;14(5):e140 [FREE Full text] [doi: 10.2196/jmir.2067] [Medline: 23111097]
18. Braun V, Clarke V. Thematic Analysis. In: Cooper HM, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology. Vol. 2. Research designs: Quantitative, qualitative neuropsychological and biological. Washington, DC: American Psychological Association (APA); 2012:57-71.
19. Kelders SM, Pots WT, Oskam MJ, Bohlmeijer ET, van Gemert-Pijnen JE. Development of a web-based intervention for the indicated prevention of depression. BMC Med Inform Decis Mak 2013;13:26 [FREE Full text] [doi: 10.1186/1472-6947-13-26] [Medline: 23425322]
20. DeLone WH, McLean ER. The DeLone and McLean Model of Information Systems Success: a Ten-Year Update. Manag Informat Syst 2003;19:9-30.
21. Patton MQ. Qualitative evaluation and research methods. 2nd ed. Newbury Park, CA: Sage Publications; 1990.
22. Mitchell AJ, Lord K, Slattery J, Grainger L, Symonds P. How feasible is implementation of distress screening by cancer clinicians in routine clinical care? Cancer 2012 Dec 15;118(24):6260-6269 [FREE Full text] [doi: 10.1002/cncr.27648] [Medline: 22674666]
23. Bennion AE, Molassiotis A. Qualitative research into the symptom experiences of adult cancer patients after treatments: a systematic review and meta-synthesis. Support Care Cancer 2013 Jan;21(1):9-25. [doi: 10.1007/s00520-012-1573-x] [Medline: 22972487]
24. Akkermans M. CBS Statistics Netherlands. 2015. Rapid increase internet usage among older people URL: http://www.cbs.nl/en-GB/menu/themas/vrije-tijd-cultuur/publicaties/artikelen/archief/2013/2013-4005-wm.htm [accessed 2015-06-26] [WebCite Cache ID 6ZZQE7752]
25. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
26. Watson M, Dunn J, Holland JC. Review of the history and development in the field of psychosocial oncology. Int Rev Psychiatry 2014 Feb;26(1):128-135. [doi: 10.3109/09540261.2013.849230] [Medline: 24716506]
27. Wagner EH, Ludman EJ, Aiello Bowles EJ, Penfold R, Reid RJ, Rutter CM, et al. Nurse navigators in early cancer care: a randomized, controlled trial. J Clin Oncol 2014 Jan 1;32(1):12-18 [FREE Full text] [doi: 10.1200/JCO.2013.51.7359] [Medline: 24276777]
28. Schwartz B. The paradox of choice: why more is less. New York, NY: Harper Collins Publishers Inc; 2004.
29. Wiggers AM, Vosbergen S, Kraaijenhagen R, Jaspers M, Peek N. Changes in the cardiac rehabilitation workflow process needed for the implementation of a self-management system. Stud Health Technol Inform 2013;192:1140. [Medline: 23920914]

Abbreviations

CW: cognitive walkthrough
HNC: head and neck cancer
PROs: patient reported outcomes
QOL: quality of life

Edited by P Bamidis; submitted 26.06.15; peer-reviewed by I Frederix, A Whitton, L van Velsen; comments to author 20.07.15; revised version received 29.08.15; accepted 22.09.15; published 21.10.15.

Please cite as:
Duman-Lubberding S, van Uden-Kraan CF, Peek N, Cuijpers P, Leemans CR, Verdonck-de Leeuw IM
An eHealth Application in Head and Neck Cancer Survivorship Care: Health Care Professionals' Perspectives
J Med Internet Res 2015;17(10):e235
URL: http://www.jmir.org/2015/10/e235/
doi:10.2196/jmir.4870
PMID:26489918

Original Paper

Text Message and Internet Support for Coronary Heart Disease Self-Management: Results From the Text4Heart Randomized Controlled Trial

Leila Pfaeffli Dale[1], MA; Robyn Whittaker[1], PhD; Yannan Jiang[1], PhD; Ralph Stewart[2], PhD; Anna Rolleston[3], PhD; Ralph Maddison[1], PhD

[1]National Institute for Health Innovation, University of Auckland, Auckland, New Zealand

[2]Department of Cardiology, Auckland City Hospital, Auckland, New Zealand

[3]Department of Medicine, University of Auckland, Auckland, New Zealand

Corresponding Author:
Leila Pfaeffli Dale, MA
National Institute for Health Innovation
University of Auckland
School of Population Health, Tamaki Campus
261 Morrin Rd
Auckland, 1072
New Zealand
Phone: 64 9 373 7599
Fax: 64 9 373 1710
Email: l.pfaeffli@auckland.ac.nz

Abstract

Background: Mobile technology has the potential to deliver behavior change interventions (mHealth) to reduce coronary heart disease (CHD) at modest cost. Previous studies have focused on single behaviors; however, cardiac rehabilitation (CR), a component of CHD self-management, needs to address multiple risk factors.

Objective: The aim was to investigate the effectiveness of a mHealth-delivered comprehensive CR program (Text4Heart) to improve adherence to recommended lifestyle behaviors (smoking cessation, physical activity, healthy diet, and nonharmful alcohol use) in addition to usual care (traditional CR).

Methods: A 2-arm, parallel, randomized controlled trial was conducted in New Zealand adults diagnosed with CHD. Participants were recruited in-hospital and were encouraged to attend center-based CR (usual care control). In addition, the intervention group received a personalized 24-week mHealth program, framed in social cognitive theory, sent by fully automated daily short message service (SMS) text messages and a supporting website. The primary outcome was adherence to healthy lifestyle behaviors measured using a self-reported composite health behavior score (≥3) at 3 and 6 months. Secondary outcomes included clinical outcomes, medication adherence score, self-efficacy, illness perceptions, and anxiety and/or depression at 6 months. Baseline and 6-month follow-up assessments (unblinded) were conducted in person.

Results: Eligible patients (N=123) recruited from 2 large metropolitan hospitals were randomized to the intervention (n=61) or the control (n=62) group. Participants were predominantly male (100/123, 81.3%), New Zealand European (73/123, 59.3%), with a mean age of 59.5 (SD 11.1) years. A significant treatment effect in favor of the intervention was observed for the primary outcome at 3 months (AOR 2.55, 95% CI 1.12-5.84; *P*=.03), but not at 6 months (AOR 1.93, 95% CI 0.83-4.53; *P*=.13). The intervention group reported significantly greater medication adherence score (mean difference: 0.58, 95% CI 0.19-0.97; *P*=.004). The majority of intervention participants reported reading all their text messages (52/61, 85%). The number of visits to the website per person ranged from zero to 100 (median 3) over the 6-month intervention period.

Conclusions: A mHealth CR intervention plus usual care showed a positive effect on adherence to multiple lifestyle behavior changes at 3 months in New Zealand adults with CHD compared to usual care alone. The effect was not sustained to the end of the 6-month intervention. A larger study is needed to determine the size of the effect in the longer term and whether the change in behavior reduces adverse cardiovascular events.

Trial Registration: ACTRN 12613000901707; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=364758&isReview=true (Archived by WebCite at http://www.webcitation.org/6c4qhcHKt)

(J Med Internet Res 2015;17(10):e237) doi:10.2196/jmir.4944

KEYWORDS

text messaging; mHealth; cellular phone; cardiovascular diseases; intervention; lifestyle change; behavior

Introduction

Coronary heart disease (CHD) remains a leading cause of death [1] and an economic burden worldwide [2]. Approximately 80% of CHD is caused by modifiable risk factors, including physical inactivity, smoking, unhealthy diet, and harmful alcohol consumption [1]. Implementing lifestyle changes and adhering to prescribed medication regimens can reduce the risk of future cardiac events and aid recovery [1]. Cardiac rehabilitation (CR) is an essential part of the contemporary management of CHD [3,4] and typically involves a program of medication and risk factor education, supervised exercise training, and psychological support.

Recent meta-analyses of randomized controlled trials (RCTs) have reported that CR is associated with improvements in mortality and morbidity [5-7], favorable cholesterol profiles [6,7], changes in smoking prevalence and blood pressure [6], and positive effects on quality of life [5,7]. Despite the benefits of CR, participation rates are less than 50% in most high- and middle-income countries [8], including New Zealand [9]. In the United States, an audit of 267,427 Medicare beneficiaries found that only 18.7% of eligible patients attended one or more outpatient CR sessions [10] and only 3 of 28 European countries estimated that CR participation was greater than 50% [11].

Most CR programs are delivered face-to-face in group sessions at hospitals or community centers. Low attendance rates indicate that the center-based approach does not suit all patients [12]. Home-based CR programs have been shown to be equally effective in clinical and health-related quality of life outcomes; however, few CR programs offer a home-based alternative [12]. Telehealth CR interventions have also shown effective reductions in CHD risk factors [13]. Greater choice of delivery model could improve CR attendance. Another option to explore is mobile CR (mHealth) because mobile technology continues to be integrated into daily life and usage rates are near 100% globally [14].

Increasingly, mHealth, the use of mobile technology (eg, short message service [SMS] text messaging, video messaging, instant messaging, the Internet, apps, and voice calling) to deliver health care, is being utilized in disease prevention and management [15,16]. Text messaging, the most researched form of mHealth, can facilitate health behavior change because it allows instant and individualized health communication and reinforcement through periodic prompts and reminders [15-18]. Recently, SMS interventions have successfully improved physical activity levels [19] and medication adherence among the CHD population [20,21]; however; CR involves supporting people to make multiple lifestyle changes because many patients have more than one behavioral risk factor. There is potential to improve individuals' overall health and reduce health care costs by targeting multiple health behaviors [22].

The aim of this study was to investigate the effectiveness of an mHealth-delivered comprehensive CR program (Text4Heart) to improve adherence to recommended lifestyle behaviors, in addition to usual care, in adults with CHD. We hypothesized that participants receiving the mHealth program would have greater adherence to lifestyle behaviors after the intervention compared to usual care alone. Secondary objectives included exploring the effects of the intervention on cardiovascular disease (CVD) risk, illness perceptions, medication adherence, self-efficacy, and anxiety and/or depression.

Methods

Design

We conducted a 6-month, 2-arm, parallel RCT in 123 adults diagnosed with CHD. The study received ethical approval (New Zealand Health and Disability Ethics Committee 13/NTA/06) and the protocol was registered and published before the conclusion of recruitment (ACTRN 12613000901707) [23]. The trial was developed and reported according to the CONSORT-EHEALTH statement (Multimedia Appendix 1).

Participants

We recruited participants from 2 large metropolitan hospitals in Auckland, New Zealand. A trained researcher screened and approached eligible patients about the study before discharge from hospital after their cardiac event. Included participants were English-speaking adults with a documented diagnosis of CHD (myocardial infarction, angina, or revascularization). Although participants were not required to have computer or Internet literacy, access to the Internet (eg, at home, work, or library) was a requirement. Participants need not own a mobile phone with text messaging capability because phones were supplied for the duration of the study if necessary. Those with untreated ventricular tachycardia, severe heart failure, life-threatening coexisting disease with life expectancy less than 1 year, and/or significant exercise limitations for reasons other than CHD were excluded.

Procedures

Eligible participants provided informed consent (see Multimedia Appendix 2) and completed a face-to-face baseline assessment in hospital, a clinic, or home setting within 4 weeks of hospital discharge. All participants received usual care, which included inpatient rehabilitation and encouragement to attend center-based CR. Traditional CR offered at the hospital recruiting sites in this study consisted of one 1-hour outpatient

education program per week for 6 weeks at a hospital or community center covering a range of topics, including cardiovascular risk factors, lifestyle change, and psychosocial support. Patients also were encouraged to attend a 16-session supervised exercise program at the participating hospital or outpatient center. Participants could take part in usual care CR from point of discharge to 6 months after their heart event. In addition to usual care, the intervention group received a 24-week mHealth program sent by automated daily text messages and access to a supporting website commencing within a week of the baseline assessment. All participants were telephoned at 3-months postrandomization to collect primary outcome data. No telephone coaching was done during this follow-up call. At 6-months postrandomization, participants were seen at a clinic or in a home setting for final follow-up assessment.

Intervention

We created and refined the Text4Heart intervention through formative and pretesting studies following the mHealth Development and Evaluation Framework [24]. Full details of the intervention, including example text messages and screenshots of the website, can be found in the published protocol [23]. In short, a theoretically framed comprehensive program of evidence-based CR guidelines [4,25,26] was delivered by text message and a supporting website over 24 weeks. The aim was to mirror current CR programs in educating patients about their cardiovascular risk factors and supporting them to make relevant lifestyle changes. Recommended lifestyle changes included stopping smoking, limiting alcohol consumption to less than 14 units of alcohol per week, eating 5 servings of fruit and vegetables per day while decreasing salt and saturated fat content, and starting and/or maintaining regular physical activity (150 minutes of moderate-to-vigorous intensity physical activity per week).

To encourage lifestyle change, the intervention was based on social cognitive theory and the key mediator of self-efficacy. Perceived self-efficacy is the extent to which people believe they can exercise control over their health behaviors [27]. With higher levels of self-efficacy, individuals can self-regulate their behavior by setting goals, creating incentives, and enlisting social support from others to maintain their motivation [28]. Self-efficacy has been shown to decline among CR nonattenders [29], which is noteworthy because higher levels of self-efficacy are linked to better clinical outcomes such as lower blood pressure and reduced hospitalizations [30].

Although targeting self-efficacy can help change lifestyle behaviors, people with CHD may need additional support to cope cognitively and emotionally with their heart event. People diagnosed with CHD can experience many negative emotions, including anxiety and depression, which can negatively impact their recovery process [31]. One way to improve coping is to modify illness perceptions and emotional representations of the disease. The Common Sense Model was also used to frame the intervention because it specifically outlines coping strategies for modifying illness perceptions and the negative emotions that arise with a health threat [32].

Participants received 7 messages per week (1 per day) and had access to a supporting website. Intervention participants also received a pedometer to self-monitor their physical activity. Messages were tailored to participants' name and preferred time of day to receive messages. From weeks 13 to 24, the frequency of messages decreased to 5 per week. Bidirectional messaging was used because participants were prompted to text in their weekly pedometer step counts and to ask questions or for feedback on other behaviors. Reponses to step counts were automated and based on the number of steps achieved, whereas individual questions were responded to personally by the research team within 48 hours. Participants were reimbursed for any costs associated with text messaging. The supporting website was accessed using a secure log-in system and included additional information, biweekly tips from the research team via a participant blog, graphs displaying their pedometer step counts, and short video messages from role models and medical professionals [33]. No changes were made to the intervention content or delivery during the study period. All text messages were sent from a centralized server.

Outcome Measures

The primary outcome was patient adherence to recommended health guidelines measured as a binary variable using a self-reported composite health behavior score based on the European Prospective Investigation into Cancer (EPIC)-Norfolk Prospective Population Study [34] at 6 months. We amended the study protocol shortly after recruitment began to include an additional end point: the same composite measure was used to collect the primary outcome at 3-months postrandomization during planned telephone calls because we decided it would be interesting to measure behavior change at the halfway point of the study in addition to 6 months. Participants received a score from 0 to 4 (out of 4) based on the number of health guidelines they met. Based on their score, participants were classified as adherent if they scored 3 or more out of 4 and nonadherent if they scored 2 or less. The health behaviors, scores, and outcome measures were smoking habit (1=not currently smoking; 0=had ≥1 cigarettes in past 7 days) as measured by a smoking history questionnaire [35], fruit and vegetable intake (1 indicates ≥5 servings daily; 0 indicates ≤4 servings daily) from the New Zealand Health Survey [36], alcohol intake (1 indicates ≤13 units per week; 0 indicates ≥14 units per week) as measured by the Alcohol Use Disorders Identification Test Consumption (AUDIT C) [37], and physical activity (1 indicates ≥14 units of moderate-to-vigorous activity/week; 0 indicates ≤13 units of moderate-to-vigorous activity/week) as measured by the Godin Leisure Time Physical Activity Questionnaire [38].

Secondary outcomes were evaluated at 6 months using self-completed questionnaires and clinical assessments. Clinical outcomes included individual biomedical risk factors (systolic and diastolic blood pressure, lipid profile, weight, body mass index, waist-to-hip ratio) and subsequent CHD risk probability using models proposed by D'Agostino developed from the Framingham Heart Study [39]. Medication adherence was measured using the Morisky 8-item Medication Adherence Questionnaire [40]. Psychological measures included the Self-efficacy for Managing Chronic Disease 6-item scale [41], the Brief Illness Perception Questionnaire [42], and the Hospital Anxiety and Depression Scale [43]. Serious adverse event data were collected at the 6-month assessment. Fidelity to the

Text4Heart intervention was assessed using an author-derived questionnaire and calculating website and response text message usage statistics (intervention group only).

Randomization and Blinding

Following informed consent and the baseline assessment, participants were randomized to either the intervention or the control group in a one-to-one ratio and stratified according to smoking status (smoker vs nonsmoker) to balance baseline health behavior scores. The randomization sequence was computer generated by a statistician independent to the project using a block size of 6. Allocation was concealed in sequentially numbered, opaque, sealed envelopes. Participant enrollment and assignment to the intervention were completed by a trained research assistant after baseline data collection. Because of the nature of the intervention, participants and outcome assessors were not blinded to their treatment allocation. Investigators, project statisticians, and usual care CR program leaders were blinded to group allocation.

Analysis

Sample Size

We estimated that a sample size of 120 (60 per group) would provide at least 80% power at the 5% level of significance (2-sided) to detect an absolute difference of 25% between the 2 groups, in the proportions of participants adherent to recommended healthy behavior guidelines. A previous study using a similar health behavior score found that approximately 70% of adults without known cardiovascular disease adhered to 3 to 4 out of 4 health behaviors [34]. We estimated that 30% of our study population with established CHD would be adherent at baseline and hypothesized that the Text4Heart intervention would change the proportion of participants' adherent to recommended healthy behavior guidelines by at least 25% compared to the control group at 6 months postrandomization.

Statistical Methods

We analyzed treatment evaluations by intention to treat, using the observed data collected from all randomized participants. Missing data were not imputed if the proportion of missing in the primary outcome was less than 10%. We did all statistical analyses using SAS version 9.3 (SAS Institute, Cary, NC, USA). All statistical tests were 2-sided at a 5% significance level. A formal statistical analysis plan was approved by the trial steering committee before data lock.

We used logistic regression to measure the main treatment effect on the proportion of participants adherent to lifestyle change (≥3 of 4 behaviors) at the end of the 6-month intervention period, adjusting for baseline adherence level and stratification factor (smoking status) and on the same outcome at 3 months. We used analysis of covariance (ANCOVA) regression to evaluate the treatment effect on continuous secondary outcomes, adjusting for baseline outcome value (if measured) and smoking status. We completed frequency and descriptive statistics on the intervention feedback survey and website and response text message usage statistics using Microsoft Excel 2010. All analyses on secondary outcomes were exploratory. We did not consider any sensitivity or subgroup analyses.

Results

Figure 1 presents the flow diagram of the progress through the phases of the trial. A total of 291 patients were screened and recruited over 10 months between 2013 and 2014. Of these, 123 eligible participants were randomized to the intervention (n=61) or the control (n=62) group.

Participants were predominantly male (100/123, 81.3%), New Zealand European (73/123, 59.3%), with a mean age of 59.5 (SD 11.1) years (see Table 1 for demographics). One quarter of participants had a household income of less than the average yearly income of NZ $50,000 (31/123, 25.2%) [44]. By the 6-month assessment, approximately half had attended at least one session of usual care CR (intervention: 30/61, 49%; control: 34/62, 55%). All participants in the intervention group used their own mobile phone.

Table 1. Participant baseline demographic and clinical characteristics (N=123).

Characteristic	Intervention (n=61)	Control (n=62)
Age (years), mean (SD)	59.0 (10.5)	59.9 (11.8)
Gender, n (%)		
Male	48 (79)	52 (84)
Female	13 (21)	10 (16)
Ethnicity [a], n (%)		
New Zealand or other European	46 (75)	45 (73)
Māori (indigenous)	6 (10)	2 (3)
Pacific Island	5 (8)	2 (3)
Indian	6 (10)	8 (13)
Other	2 (3)	5 (8)
Income (NZ$) [b], n (%)		
<50,000/year	14 (23)	17 (27)
>50,000/year	39 (64)	40 (65)
Don't know/refuse to answer	8 (13)	5 (8)
Cardiac diagnosis, n (%)		
Myocardial infarction	46 (75)	52 (84)
Unstable angina	4 (7)	5 (8)
Angina	11 (18)	5 (8)
Cardiac procedure, n (%)		
Percutaneous coronary intervention	43 (70)	47 (76)
Coronary artery bypass grafting	14 (23)	10 (16)
Medical management	4 (7)	5 (8)
Diabetes	14 (23)	7 (11)

[a]Could identify with more than 1 ethnicity.

[b]Income split into categories based on earning less or greater than the average yearly income of NZ $50,000.

For the primary outcome, the intervention group increased adherence to recommended lifestyle behavior changes from 33% (20/61) at baseline to 59% (36/61) at 3 months and then plateaued with 53% (32/61) still adherent at 6 months. The control group had a smaller increase in adherence from 27% (17/62) at baseline to 37% (23/62) at 3 months and 39% (24/62) at 6 months. A significant treatment effect in favor of the intervention was observed at 3 months (AOR 2.55, 95% CI 1.12-5.84; *P*=.03), but not at 6 months (AOR 1.93, 95% CI 0.83-4.53; *P*=.13). The percentage adherent to individual behaviors can be seen in Table 2.

For the secondary outcomes (Table 3), the intervention group reported a significantly greater medication adherence score (mean difference: 0.58, 95% CI 0.19-0.97; *P*=.004). The intervention group also had lower low-density lipoprotein (LDL) cholesterol than the control group (mean difference: –0.25, 95% CI –0.49 to 0.01; *P*=.05) at 6 months, although this did not meet statistical significance. A negative effect was seen for total hospital anxiety with the intervention group reporting significantly greater anxiety than the control group at 6 months (mean difference: 1.18, 95% CI 0.28-2.08; *P*=.01). No differences were seen for clinical or other psychological outcomes. There were 13 (intervention: n=8; control: n=5) serious adverse events reported during the trial, although none were study related.

Table 2. Adherence to individual behaviors at baseline, 3 months, and 6 months.

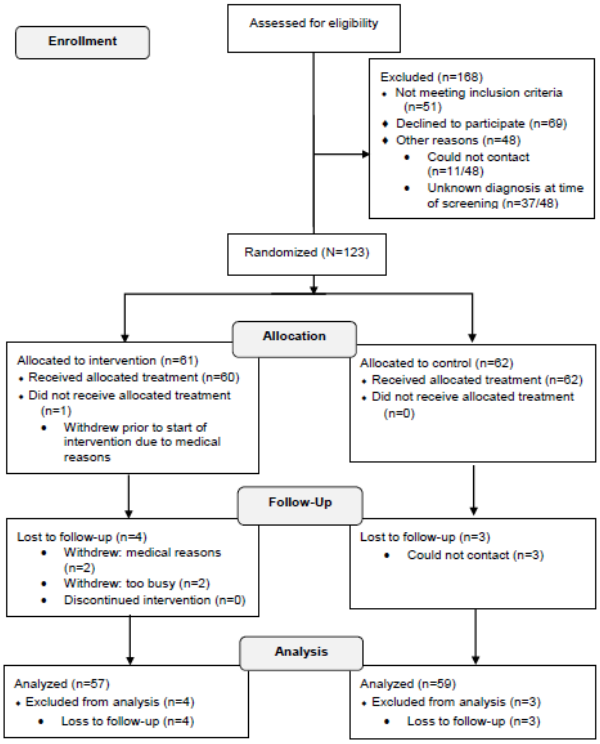
Individual behavior	Intervention, n (%) (n=61)	Control, n (%) (n=62)
Baseline		
Nonsmoker	49 (80)	51 (82)
Nonharmful alcohol intake	53 (87)	53 (86)
Physically active	17 (28)	7 (11)
≥5 Fruit and vegetable intake	12 (20)	15 (24)
3 months		
Nonsmoker	52 (85)	53 (86)
Nonharmful alcohol intake	56 (92)	53 (88)
Physically active	21 (34)	10 (16)
≥5 Fruit and vegetable intake	33 (54)	18 (29)
6 months		
Nonsmoker	51 (84)	55 (89)
Nonharmful alcohol intake	53 (87)	56 (90)
Physically active	19 (31)	15 (24)
≥5 Fruit and vegetable intake	29 (48)	15 (24)

Table 3. Baseline and 6-month secondary outcomes.

Outcome	Intervention, mean (SD) (n=61)		Control, mean (SD) (n=62)		Adjusted difference (95% CI) at 6 months	*P*
	Baseline	6 months	Baseline	6 months		
Clinical outcomes						
BMI	31.0 (6.4)	30.3 (5.4)	28 (4.2)	28.1 (4.4)	–0.10 (–0.56 to 0.35)	.66
Waist-to-hip ratio	0.98 (0.07)	0.97 (0.06)	0.95 (0.07)	0.94 (0.07)	0.01 (–0.01 to 0.02)	.29
Blood pressure (mm Hg)						
Systolic	131 (17)	136 (20)	129 (26)	135 (16)	0.09 (–6.43 to 6.61)	.98
Diastolic	78 (11)	79 (11)	75 (11)	79 (10)	–0.24 (–3.86 to 3.38)	.90
Cholesterol (mmol/L)						
Total	4.6 (1.2)	3.6 (0.7)	4.3 (1.2)	3.8 (1.1)	–0.29 (–0.61 to 0.03)	.08
HDL	1.1 (0.3)	1.1 (0.3)	1.1 (0.3)	1.2 (0.4)	–0.04 (–0.15 to 0.07)	.51
LDL	2.7 (1.3)	1.7 (0.6)	2.4 (1.0)	1.9 (0.8)	–0.25 (–0.49 to 0.01)	.053
CVD risk probability		7.9 (3.4)		8.1 (3.3)	–0.27 (–1.58 to 1.04)	.68
Medication adherence[a]		7.3 (0.9)		6.8 (1.2)	0.58 (0.19 to 0.97)	.004
Psychological outcomes						
Overall illness threat	41.8 (12.3)	32.7 (11.2)	39.8 (11.6)	32.1 (12.6)	–0.4 (–4.18 to 3.35)	.83
Hospital anxiety	6.3 (3.9)	5.8 (3.5)	5.5 (3.5)	4.4 (2.9)	1.18 (0.28 to 2.08)	.01
Hospital depression	4.3 (3.3)	2.8 (2.8)	3.8 (2.3)	2.5 (2.2)	0.08 (–0.71 to 0.87)	.84
Overall self-efficacy	7.6 (1.6)	8.1 (1.48)	7.9 (1.4)	8.3 (1.2)	–0.07 (–0.47 to 0.33)	.73

[a]Use of the MMAS is protected by US copyright laws. Permission for use is required. A license agreement is available from Donald E Morisky, ScD, ScM, MSPH, Professor, Department of Community Health Services, UCLA School of Public Health, 650 Charles E Young Drive South, Los Angeles, CA 90095-1772, United States.

Figure 1. Trial registration flowchart.



Intervention Fidelity and Acceptability

All but one participant randomized to the intervention group received the Text4Heart program. Intervention participants reported high fidelity to the text messaging component; 52 of 61 (85%) participants reported reading all their text messages. Nearly all participants sent in at least one step count text response (58/61, 95%) with a mean of 15 (SD 8.7) step count replies per participant over 24 weeks. A total of 23 participants sent in questions or comments to the study team via text (23/61, 38%). The vast majority of participants (55/61, 90%) felt using text messaging was a good way to deliver the Text4Heart program. Most felt that the 24-week program was the right length (48/61, 79%) and that we sent the right number of text messages (51/61, 84%). Only 5 of 61 participants (8%) felt we sent too many messages.

Less than half of participants (26/61, 43%) felt using a website was a good way to deliver the Text4Heart program. Website use data showed that 75% of participants (46/61) logged onto the website at least once during the intervention period. The number of visits to the website per person ranged from 0 to 100 (median 3) over the 6-month intervention period. Two participants reported not using the website because they did not know how to use it. Despite the lack of website use, nearly all participants (55/61, 90%) would recommend the Text4Heart program (both text message and Web) to other people who have had a heart event. Most participants felt the program helped them learn about (47/61, 77%) and recover (51/61, 84%) from their heart event. More detail on participants' perceptions about Text4Heart can be found in Multimedia Appendix 3.

Discussion

The Text4Heart intervention improved adherence to lifestyle behaviors at 3 months when compared to usual care alone (control), although the size of effect was not significantly retained at 6 months. A treatment effect was also observed for medication adherence. This study is one of the first to demonstrate a positive effect of an SMS-based intervention on multiple lifestyle behaviors. These findings highlight the potential utility of this approach to augment existing services in people with CHD. Strengths of the study included the RCT design, minimal loss to follow-up, high fidelity to the text messaging feature of the intervention, and the use of a composite health behavior score. A composite score allowed for a clear understanding of the intervention's overall impact without the risk of increasing type 1 error rate [22,45].

The Text4Heart intervention was delivered in addition to usual care, which included inpatient CR and encouragement to attend phase II outpatient CR. Because both groups had similar phase II CR attendance, it appeared receiving simple text messages resulted in greater lifestyle change at 3 months. This study was powered to detect a large effect (25% difference between groups) at 6 months, yet the observed difference at this end point was 14%. A small improvement in adherence to multiple lifestyle behaviors may still have clinical relevance. To detect an effect of this magnitude (14%-15%), a post hoc sample size calculation indicated that 400 participants would be needed; thus, a larger study is warranted to examine any sustained effects of such an intervention.

Small changes to the Text4Heart program may also help boost the effects from 3 to 6 months. Relapse in unhealthy behaviors may have occurred when the intensity of our text messaging

decreased. Relapse prevention and coping content should be delivered to re-engage those who drop off. Relapse prevention has been investigated in smoking cessation trials [46]; however, few studies have reported on maintenance of other healthy behaviors and strategies to prevent relapse [47] and this remains an area of future research.

Limitations

Although treatment allocation was concealed before randomization, a limitation of this trial was that the outcome assessors were not blinded (participants were randomized at the conclusion of baseline visits by the same outcome assessors who conducted follow-up visits at 6 months). In addition, the primary outcome measure was self-reported so recall bias is possible, although validated questionnaires were used where feasible. The composite score did not capture all aspects of behavior; however, we felt it was appropriate because it is difficult to measure the multiple outcomes of CR. Objective clinical measures provided additional information associated with behavior changes; however, due to the short follow-up these findings were exploratory in nature. Another limitation was that the findings may not be transferable to other populations because our sample was predominantly New Zealand European, earned higher than the average yearly income, and were generally text message and computer literate.

Comparisons With Other Work

Our findings extend previous research supporting the use of text message-delivered interventions to promote behavior change in people with CHD [19,48]. Text4Heart is one of the few to intervene on and measure multiple behaviors, mirroring traditional CR, which focuses on all potential behavioral risk factors. Intervening on multiple behaviors simultaneously has the potential to maximize the impact on an individual's health and may also be more cost-effective than addressing one behavior at a time; however, it is unknown whether changing behaviors sequentially is more effective than simultaneously because few studies have compared the 2 approaches [22].

We found a positive treatment effect on medication adherence supporting other text message interventions shown to improve antiplatelet medication adherence among CHD patients [21]. The Text4Heart program incorporated several essential intervention components needed for improved medication adherence, namely, patient knowledge, counseling, and self-monitoring [49]. Future iterations might encourage greater patient counseling because the 2-way text message communication option in this study was underutilized, which may lead to stronger outcomes. The greater medication adherence score and the increased servings of fruit and vegetables observed (a component of the health behavior score) may have contributed to lower LDL cholesterol among the intervention group at 6 months. A larger trial with longer follow-up is needed to determine whether the Text4Heart intervention can make a clinically significant difference on LDL cholesterol levels.

An unexpected outcome was the lower level of anxiety observed in the control group. Both groups' mean anxiety score was in the normal category (0-7) and decreased from baseline to 6 months. A possible explanation for the effect on anxiety might be that receiving text messages about one's disease may lead to more anxiety. Previous research has shown that confronting a CHD diagnosis elicited negative emotions in the short term, but improved health in the long term [50]. Longer follow-up is needed to see if the difference observed in Text4Heart persisted over time.

No differences were found between groups in the self-efficacy or illness perception constructs. Both groups had high scores at baseline, which may have led to a ceiling effect. Future studies should focus on patients who have low self-efficacy to change behavior.

Implications for Clinical Practice and/or Research

The Text4Heart intervention was a simple package of text messaging and a website; however, because the website was used infrequently and usual care CR attendance was similar across groups, receiving text messages was likely the predominant contributor to the observed effects. A text messaging program could be easily incorporated into existing CR, either as an alternative option for those unable to attend center-based programs or as an add-on to extend current services. It may be that earlier program commencement after diagnosis and the longer duration of Text4Heart helped facilitate and maintain behavior change. There is potential for this type of program to reach underserved populations and those with access-to-care barriers, such as patients living in deprived areas or developing nations. Future research should be undertaken with more diverse samples to determine if such an intervention can reduce health inequalities.

Text4Heart was relatively simple to develop and use. A similar study delivering exercise-based CR via text message was considered to be cost-effective for walking and leisure-time physical activity [19]. Apps [48] and the use of biofeedback and wearable sensors have also been used to deliver CR. Apps and sensor technologies may result in stronger effects because they can allow for greater individual tailoring and improved 2-way communication with health care providers; however, such technology are associated with greater financial and time costs. More research is needed to compare outcomes of text messaging and "appified" approaches before investing significant resources into complex interventions.

Conclusion

Receiving a simple text message-delivered CR intervention in addition to usual care had a positive effect on adherence to multiple lifestyle behavior changes in New Zealand adults with CHD at 3 months; however, the effect had attenuated by 6 months. A larger study with longer follow-up is needed to determine whether these behavior changes can result in clinically significant outcomes.

Acknowledgments

We wish to acknowledge the contributions of the cardiac rehabilitation nurses in recruitment and to the content of the Text4Heart intervention; in particular, Adele Clayton. We would also like to acknowledge the National Institute for Health Innovation's IT team for their work on the content delivery system. The study was funded in part by a Health Research Council Sir Charles Hercus Fellowship and a HOPE Selwyn Foundation Scholarship in Ageing Research. Dr Maddison was supported by a Health Research Council Sir Charles Hercus Fellowship.

Conflicts of Interest

None declared.

Multimedia Appendix 1

eHealth CONSORT checklist form.

[PDF File (Adobe PDF File), 7MB - jmir_v17i10e237_app1.pdf]

Multimedia Appendix 2

Participant information sheet and informed consent form.

[PDF File (Adobe PDF File), 95KB - jmir_v17i10e237_app2.pdf]

Multimedia Appendix 3

Participant perceptions of the Text4Heart intervention.

[PDF File (Adobe PDF File), 4KB - jmir_v17i10e237_app3.pdf]

References

1. World Health Organization. 2015. Fact sheet N317: cardiovascular diseases (CVDs) URL: http://www.who.int/mediacentre/factsheets/fs317/en/ [accessed 2015-07-16] [WebCite Cache ID 6a4pZNElL]
2. World Health Organization, World Economic Forum. Essential Medicines and Health Products Information Portal. 2011. From burden to best buys: reducing the economic impact of NCDs in low- and middle-income countries URL: http://apps.who.int/medicinedocs/en/d/Js18804en/ [accessed 2015-07-16] [WebCite Cache ID 6a4sGyrXI]
3. Balady GJ, Williams MA, Ades PA, Bittner V, Comoss P, Foody JM, American Heart Association Exercise, Cardiac Rehabilitation,Prevention Committee, the Council on Clinical Cardiology, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on EpidemiologyPrevention, American Heart Association Council on Nutrition, Physical Activity,Metabolism, American Association of Cardiovascular and Pulmonary Rehabilitation. Core components of cardiac rehabilitation/secondary prevention programs: 2007 update: a scientific statement from the American Heart Association Exercise, Cardiac Rehabilitation, and Prevention Committee, the Council on Clinical Cardiology; the Councils on Cardiovascular Nursing, Epidemiology and Prevention, and Nutrition, Physical Activity, and Metabolism; and the American Association of Cardiovascular and Pulmonary Rehabilitation. Circulation 2007 May 22;115(20):2675-2682 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.106.180945] [Medline: 17513578]
4. Piepoli M, Corrà U, Adamopoulos S, Benzer W, Bjarnason-Wehrens B, Cupples M, et al. Secondary prevention in the clinical management of patients with cardiovascular diseases. Core components, standards and outcome measures for referral and delivery: a policy statement from the cardiac rehabilitation section of the European Association for Cardiovascular Prevention & Rehabilitation. Endorsed by the Committee for Practice Guidelines of the European Society of Cardiology. Eur J Prev Cardiol 2014 Jun;21(6):664-681. [doi: 10.1177/2047487312449597] [Medline: 22718797]
5. Heran B, Chen J, 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]
6. Lawler P, Filion K, Eisenberg M. 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.e2. [doi: 10.1016/j.ahj.2011.07.017] [Medline: 21982647]
7. Clark A, Hartling L, Vandermeer B, McAlister FA. Meta-analysis: secondary prevention programs for patients with coronary artery disease. Ann Intern Med 2005 Nov 1;143(9):659-672. [Medline: 16263889]
8. Turk-Adawi K, Grace S. Narrative review comparing the benefits of and participation in cardiac rehabilitation in high-, middle- and low-income countries. Heart Lung Circ 2015 May;24(5):510-520 [FREE Full text] [doi: 10.1016/j.hlc.2014.11.013] [Medline: 25534902]

9. Doolan-Noble F, Broad J, Riddell T, North D. Cardiac rehabilitation services in New Zealand: access and utilisation. N Z Med J 2004 Jul 9;117(1197):U955. [Medline: 15326508]
10. Suaya J, Shepard D, Normand SLT, Ades PA, Prottas J, Stason WB. Use of cardiac rehabilitation by Medicare beneficiaries after myocardial infarction or coronary bypass surgery. Circulation 2007 Oct 9;116(15):1653-1662 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.701466] [Medline: 17893274]
11. Bjarnason-Wehrens B, McGee H, Zwisler A, Piepoli MF, Benzer W, Schmid JP, Cardiac Rehabilitation Section European Association of Cardiovascular PreventionRehabilitation. Cardiac rehabilitation in Europe: results from the European Cardiac Rehabilitation Inventory Survey. Eur J Cardiovasc Prev Rehabil 2010 Aug;17(4):410-418. [doi: 10.1097/HJR.0b013e328334f42d] [Medline: 20300001]
12. Dalal H, Zawada A, Jolly K, Moxham T, Taylor RS. Home based versus centre based cardiac rehabilitation: Cochrane systematic review and meta-analysis. BMJ 2010;340:b5631 [FREE Full text] [Medline: 20085991]
13. Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review. Eur J Cardiovasc Prev Rehabil 2009 Jun;16(3):281-289. [doi: 10.1097/HJR.0b013e32832a4e7a] [Medline: 19407659]
14. The World in 2013: ICT Facts and Figures. Geneva: International Telecommunication Union; 2013. URL: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf [accessed 2015-07-16] [WebCite Cache ID 6a4qDp6DT]
15. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
16. Fjeldsoe B, Marshall A, Miller Y. 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]
17. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. J Telemed Telecare 2011;17(1):41-48. [doi: 10.1258/jtt.2010.100322] [Medline: 21097565]
18. Fry J, Neff R. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
19. Maddison R, Pfaeffli L, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
20. Quilici J, Fugon L, Beguin S, Morange PE, Bonnet JL, Alessi MC, et al. Effect of motivational mobile phone short message service on aspirin adherence after coronary stenting for acute coronary syndrome. Int J Cardiol 2013 Sep 20;168(1):568-569. [doi: 10.1016/j.ijcard.2013.01.252] [Medline: 23462636]
21. Park L, Howie-Esquivel J, Chung M, Dracup K. A text messaging intervention to promote medication adherence for patients with coronary heart disease: a randomized controlled trial. Patient Educ Couns 2014 Feb;94(2):261-268. [doi: 10.1016/j.pec.2013.10.027] [Medline: 24321403]
22. Prochaska J, Spring B, Nigg C. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
23. Dale LP, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Improving coronary heart disease self-management using mobile technologies (Text4Heart): a randomised controlled trial protocol. Trials 2014;15:71 [FREE Full text] [doi: 10.1186/1745-6215-15-71] [Medline: 24588893]
24. Dale LP, Whittaker R, Eyles H, Ni Mhurchu C, Ball K, Smith N, et al. Cardiovascular disease self-management: pilot testing of an mHealth healthy eating program. J Pers Med 2014;4(1):88-101 [FREE Full text] [doi: 10.3390/jpm4010088] [Medline: 25562145]
25. New Zealand Guidelines Group. 2002. Evidence-based best practice guideline: cardiac rehabilitation URL: http://www.health.govt.nz/publication/cardiac-rehabilitation-guideline [accessed 2015-07-14] [WebCite Cache ID 6a4qzKmTv]
26. American College of Sports Medicine. ACSM'S Guidelines for Exercise Testing and Prescription. Baltimore: Lippincott; 2010.
27. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
28. Bandura A. Self-efficacy in health functioning. In: Ayers S, Baum A, McManus C, editors. Cambridge Handbook of Psychology, Health and Medicine. Cambridge: Cambridge University Press; 2007.
29. Sharp P, Salyer J. Self-efficacy and barriers to healthy diet in cardiac rehabilitation participants and nonparticipants. J Cardiovasc Nurs 2012;27(3):253-262. [doi: 10.1097/JCN.0b013e31821efdc2] [Medline: 21760521]
30. Katch H, Mead H. The role of self-efficacy in cardiovascular disease self-management: a review of effective programs. Patient Intelligence 2010 Oct;2:33-44. [doi: 10.2147/PI.S12624]
31. Hare D, Toukhsati S, Johansson P, Jaarsma T. Depression and cardiovascular disease: a clinical review. Eur Heart J 2014 Jun 1;35(21):1365-1372 [FREE Full text] [doi: 10.1093/eurheartj/eht462] [Medline: 24282187]
32. Cameron L, Jago L. Emotion regulation interventions: a common-sense model approach. Br J Health Psychol 2008;13(Pt 2):215-221. [doi: 10.1348/135910708X288800] [Medline: 18302809]
33. Text4Heart. URL: https://text4heart.co.nz/ [accessed 2015-07-17] [WebCite Cache ID 6a4nip04v]

34. Khaw K, Wareham N, Bingham S, Welch A, Luben R, Day N. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. PLoS Med 2008 Jan 8;5(1):e12 [FREE Full text] [doi: 10.1371/journal.pmed.0050012] [Medline: 18184033]
35. Maddison R, Roberts V, Bullen C, McRobbie H, Jiang Y, Prapavessis H, et al. Design and conduct of a pragmatic randomized controlled trial to enhance smoking-cessation outcomes with exercise: The Fit2Quit study. Mental Health and Physical Activity 2010 Dec;3(2):92-101. [doi: 10.1016/j.mhpa.2010.09.003]
36. Wellington: Ministry of Health. 2008. New Zealand Health Survey 2006/07 Adult Questionnaire URL: http://www.health.govt.nz/publication/portrait-health-key-results-2006-07-new-zealand-health-survey [accessed 2015-07-15] [WebCite Cache ID 6a4rUHfGO]
37. Bush K, Kivlahan D, McDonell M, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [Medline: 9738608]
38. Godin G, Shephard R. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
39. D'Agostino R, Russell M, Huse D, Ellison RC, Silbershatz H, Wilson PW, et al. Primary and subsequent coronary risk appraisal: new results from the Framingham study. Am Heart J 2000 Feb;139(2 Pt 1):272-281. [Medline: 10650300]
40. Morisky D, Ang A, Krousel-Wood M, Ward HJ. Predictive validity of a medication adherence measure in an outpatient setting. J Clin Hypertens (Greenwich) 2008 May;10(5):348-354 [FREE Full text] [Medline: 18453793]
41. Lorig K, Sobel D, Ritter P, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262 [FREE Full text] [Medline: 11769298]
42. Broadbent E, Petrie K, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res 2006 Jun;60(6):631-637. [doi: 10.1016/j.jpsychores.2005.10.020] [Medline: 16731240]
43. Snaith R. The hospital anxiety and depression scale. Health Qual Life Outcomes 2003;1:29 [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
44. MacPherson L. Statistics New Zealand. 2014 Oct 03. New Zealand Income Survey: June 2014 quarter URL: http://www.stats.govt.nz/browse_for_stats/income-and-work/Income/NZIncomeSurvey_HOTPJun14qtr.aspx [accessed 2015-07-16] [WebCite Cache ID 6a4rqLUNI]
45. Drake B, Quintiliani L, Sapp A, Li Y, Harley AE, Emmons KM, et al. Comparing strategies to assess multiple behavior change in behavioral intervention studies. Transl Behav Med 2013 Mar 1;3(1):114-121 [FREE Full text] [doi: 10.1007/s13142-013-0195-2] [Medline: 23504621]
46. Hajek P, Stead L, West R, Jarvis M, Hartmann-Boyce J, Lancaster T. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2013;8:CD003999. [doi: 10.1002/14651858.CD003999.pub4] [Medline: 23963584]
47. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol 2011 Jan;30(1):99-109. [doi: 10.1037/a0021974] [Medline: 21299298]
48. Varnfield M, Karunanithi M, Lee CK, 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]
49. Zullig L, Peterson E, Bosworth H. Ingredients of successful interventions to improve medication adherence. JAMA 2013 Dec 25;310(24):2611-2612. [doi: 10.1001/jama.2013.282818] [Medline: 24264605]
50. van Elderen T, Maes S, Dusseldorp E. Coping with coronary heart disease: a longitudinal study. J Psychosom Res 1999 Aug;47(2):175-183. [Medline: 10579500]

Abbreviations

CHD: coronary heart disease
CR: cardiac rehabilitation
LDL: low-density lipoprotein
RCT: randomized controlled trials
SMS: short message service
CVD: cardiovascular disease

Edited by G Eysenbach; submitted 16.07.15; peer-reviewed by F Doolan-Noble, M Varnfield; comments to author 02.09.15; revised version received 08.09.15; accepted 22.09.15; published 21.10.15.

Please cite as:
Pfaeffli Dale L, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R
Text Message and Internet Support for Coronary Heart Disease Self-Management: Results From the Text4Heart Randomized Controlled Trial
J Med Internet Res 2015;17(10):e237
URL: http://www.jmir.org/2015/10/e237/
doi:10.2196/jmir.4944
PMID:26490012

Original Paper

Diabetes Prevention and Weight Loss with a Fully Automated Behavioral Intervention by Email, Web, and Mobile Phone: A Randomized Controlled Trial Among Persons with Prediabetes

Gladys Block[1,2*], PhD; Kristen MJ Azar[3*], RN, MSN, MPH; Robert J Romanelli[3], MPH, PhD; Torin J Block[1], BA; Donald Hopkins[1], MS; Heather A Carpenter[1], BA; Marina S Dolginsky[3], BS; Mark L Hudes[4], PhD; Latha P Palaniappan[3], MD, MS; Clifford H Block[1], PhD

[1]Turnaround Health, a division of NutritionQuest, Berkeley, CA, United States

[2]Division of Community Health and Human Development, School of Public Health, University of California, Berkeley, Berkeley, CA, United States

[3]Palo Alto Medical Foundation Research Institute, Palo Alto, CA, United States

[4]Center for Weight and Health, University of California, Berkeley, Berkeley, CA, United States

[*]these authors contributed equally

Corresponding Author:
Gladys Block, PhD
Turnaround Health
a division of NutritionQuest
15 Shattuck Square
Suite 288
Berkeley, CA, 94704
United States
Phone: 1 510 704 8514
Fax: 1 510 704 8996
Email: gblock@berkeley.edu

Abstract

Background: One-third of US adults, 86 million people, have prediabetes. Two-thirds of adults are overweight or obese and at risk for diabetes. Effective and affordable interventions are needed that can reach these 86 million, and others at high risk, to reduce their progression to diagnosed diabetes.

Objective: The aim was to evaluate the effectiveness of a fully automated algorithm-driven behavioral intervention for diabetes prevention, Alive-PD, delivered via the Web, Internet, mobile phone, and automated phone calls.

Methods: Alive-PD provided tailored behavioral support for improvements in physical activity, eating habits, and factors such as weight loss, stress, and sleep. Weekly emails suggested small-step goals and linked to an individual Web page with tools for tracking, coaching, social support through virtual teams, competition, and health information. A mobile phone app and automated phone calls provided further support. The trial randomly assigned 339 persons to the Alive-PD intervention (n=163) or a 6-month wait-list usual-care control group (n=176). Participants were eligible if either fasting glucose or glycated hemoglobin A1c (HbA1c) was in the prediabetic range. Primary outcome measures were changes in fasting glucose and HbA1c at 6 months. Secondary outcome measures included clinic-measured changes in body weight, body mass index (BMI), waist circumference, triglyceride/high-density lipoprotein cholesterol (TG/HDL) ratio, and Framingham diabetes risk score. Analysis was by intention-to-treat.

Results: Participants' mean age was 55 (SD 8.9) years, mean BMI was 31.2 (SD 4.4) kg/m^2, and 68.7% (233/339) were male. Mean fasting glucose was in the prediabetic range (mean 109.9, SD 8.4 mg/dL), whereas the mean HbA1c was 5.6% (SD 0.3), in the normal range. In intention-to-treat analyses, Alive-PD participants achieved significantly greater reductions than controls in fasting glucose (mean –7.36 mg/dL, 95% CI –7.85 to –6.87 vs mean –2.19, 95% CI –2.64 to –1.73, P<.001), HbA1c (mean –0.26%, 95% CI –0.27 to –0.24 vs mean –0.18%, 95% CI –0.19 to –0.16, P<.001), and body weight (mean –3.26 kg, 95% CI –3.26 to –3.25 vs mean –1.26 kg, 95% CI –1.27 to –1.26, P<.001). Reductions in BMI, waist circumference, and TG/HDL were also significantly greater in Alive-PD participants than in the control group. At 6 months, the Alive-PD group reduced their Framingham 8-year diabetes risk from 16% to 11%, significantly more than the control group (P<.001). Participation and retention

was good; intervention participants interacted with the program a median of 17 (IQR 14) of 24 weeks and 71.1% (116/163) were still interacting with the program in month 6.

Conclusions: Alive-PD improved glycemic control, body weight, BMI, waist circumference, TG/HDL ratio, and diabetes risk. As a fully automated system, the program has high potential for scalability and could potentially reach many of the 86 million US adults who have prediabetes as well as other at-risk groups.

Trial Registration: Clinicaltrials.gov NCT01479062; https://clinicaltrials.gov/ct2/show/NCT01479062 (Archived by WebCite at http://www.webcitation.org/6bt4V20NR)

(J Med Internet Res 2015;17(10):e240) doi:10.2196/jmir.4897

KEYWORDS

type 2 diabetes; prevention; intervention studies; prediabetes; behavior change; obesity; physical activity; nutrition; Internet; smartphone; weight loss

Introduction

In the United States, 86 million adults have prediabetes [1], a condition characterized by elevated blood glucose that is not yet high enough to be diagnosed as diabetes. Chronic elevated blood glucose levels tend to increase over time and it is estimated that as many as 70% of those with prediabetes will eventually progress to type 2 diabetes [2]. The economic burden of the combined costs of diabetes and prediabetes exceeded US $322 billion in 2012 and accounted for 1 in 10 US health care dollars. In an editorial, Cefalu et al [3] noted that “increased prevalence, not increased cost per patient, is the driving force behind the increased economic burden of diabetes” [4]. Unless changes are made to prevent progression to type 2 diabetes, costs relating to diabetes management and care will continue to rise at alarming rates. It is critical to develop affordable and effective interventions that can reach more of the 86 million people with prediabetes with programs to improve glycemic control.

Lifestyle modification has been shown to reduce risk of progression to diabetes by as much as 40% to 70% [2]. The Diabetes Prevention Program (DPP) achieved a 58% reduction in the incidence of diabetes through increased physical activity, dietary changes, and weight loss [5]. The DPP involved intensive counseling and multiple in-person and group meetings in a research context. Since then, numerous translations of the DPP have been developed that attempt to provide approaches that can be widely applied.

Some adaptations of the DPP for real-world settings deliver the interventions through group meetings and in-person contact, such as those delivered in communities and YMCAs [6-8]. Ali et al [9] found a mean 4.3% body weight loss in programs delivered by medical professionals and 3.2% weight loss for those delivered by community members. Although in-person and group-based interventions are important and effective resources, barriers to widespread adoption of such programs include lack of professional staff, institutional resources, substantial costs, and the requirement that participants attend a series of in-person meetings, which together substantially limit their scalability and reach [10,11].

A number of interventions have combined some form of human coaching with the use of technology, at least by phone or email, thus enabling them to achieve wider reach. In a meta-analysis of programs modeled on the DPP, Ali et al [9] found that among electronic media-assisted programs, there was a statistically significant mean weight loss of 4% body weight. A review by Levine et al [12] of technology-assisted weight loss interventions in primary care found a mean weight loss in the intervention group of -2.7 kg among technology-assisted weight loss interventions that included some human coaching. Human feedback and coaching can provide value and effectiveness—and indeed is needed by some participants. However, it does result in higher costs that once again limit the number of persons with prediabetes that can be reached.

Fully automated behavioral intervention systems, those without any human coaching or facilitation, may hold substantial promise in overcoming barriers to widespread reach and adoption in a resource-limited health care environment if they can be shown to be effective. Several such programs have been found to be effective for weight loss [13,14], but there is very little information on the impact of such programs on glycemic markers critical for diabetes prevention. The Alive-PD intervention (Turnaround Health, a Division of NutritionQuest, Berkeley, CA, USA) provides such a fully automated, tailored, online behavior change program. Alive-PD is focused on reducing diabetes risk by reducing the biomarkers that constitute the criteria for diabetes, glycated hemoglobin A_{1c} (HbA_{1c}) and fasting glucose, in persons at risk of developing diabetes. The purpose of this analysis is to examine the effects of this automated program on those glycemic biomarkers and weight loss in a randomized controlled trial.

Methods

The Alive-PD study was a randomized, wait-list controlled (usual care) trial among patients with clinical evidence of prediabetes. The primary outcome measures were changes in HbA_{1c} and fasting glucose at 6-month follow-up from baseline. Secondary outcomes were changes in body weight, body mass index (BMI), waist circumference, triglyceride (TG) to high-density lipoprotein cholesterol (HDL-C) ratio (a proxy measure for insulin resistance [15]), and metabolic syndrome. Metabolic syndrome was defined as 3 or more of 5 components (ie, abdominal obesity, elevated blood pressure, elevated TG, low HDL, and dysglycemia) specified by the American Heart Association and the National Heart, Lung, and Blood Institute [16]. The Framingham 8-year diabetes risk score was calculated

[17]. Sample size was determined by using the estimated standard deviation of change in HbA_{1c} from an intervention study on patients with diabetes [18]. With a standard deviation of 1.4 and alpha of .05, we estimated that a final sample of 268 participants would provide 80% power to detect a minimum detectable difference in change in HbA_{1c} of 0.48%. The goal for enrollment was 314 persons to achieve a sample size of 268 after 15% estimated attrition. The trial design and methods are described in detail elsewhere [19] and are summarized here (see Multimedia Appendix 1 for CONSORT flow diagram).

Participant Recruitment and Eligibility Criteria

Potential participants whose recent fasting glucose and/or HbA_{1c} were within the prediabetes range were initially identified through an electronic health record query of patients in an ambulatory care health care delivery system, the Palo Alto Medical Foundation (PAMF). The PAMF is a community-based multispecialty group practice in Northern California. Patients meeting these criteria were recruited via letter and underwent telephone screening for eligibility. Those meeting preliminary criteria were invited to attend a clinic visit to confirm eligibility, which also provided the baseline data for those confirmed eligible. At that visit, fasting glucose and lipids were measured by point-of-care whole blood testing using the Alere Cholestech LDX Analyzer. Similarly, HbA_{1c} was measured using the Siemens DCA Vantage Analyzer. Biometric measurements, including height, body weight, waist circumference, and blood pressure were also obtained. BMI (kg/m^2) was calculated from height and body weight.

Individuals were eligible if they were aged between 30 and 69 years with a BMI of at least 27 kg/m^2 (BMI >25 kg/m^2 for Asian participants) [20], spoke English, were not taking diabetes medications, had access to email and Internet, and had either fasting glucose or HbA_{1c} in the prediabetes range (glucose: 5.55-6.94 mmol/L or 100-125 mg/dL; HbA_{1c}: 39-46 mmol/mol or 5.7%-6.4%). If one measure reached the diabetic range and the other was prediabetic, the patient's primary care physician decided whether the patient had prediabetes and was eligible for the study. Additional exclusion criteria are described elsewhere [19]. The study was approved by independent Institutional Review Boards of Turnaround Health and PAMF.

After participants provided signed informed consent, they were given brief (5-10 minutes) instruction that they were at risk for developing diabetes and that increased physical activity and changes in their dietary behaviors could help prevent progression to diabetes. PAMF research staff assisted participants in signing into an account for the Alive-PD Web-based program, where participants provided their email address and password to the system. All subsequent communications with participants came from the electronic Alive-PD program and interactions with the Alive-PD program took place outside of the clinic.

Randomization

After leaving the study site, enrolled participants completed a brief questionnaire online, which provided information required for randomization. Randomization was conducted automatically, by computer algorithm, with stratification by sex, race/ethnicity (non-Hispanic white/other), and BMI (<35 kg/m^2/≥35 kg/m^2), to achieve balance on those factors. Participants were randomized to start the intervention immediately (intervention group) or after 6 months' delay (control group/wait-listed usual care group). Participants were notified of treatment group assignment by automated email from the Alive-PD system. The research and clinical staff at PAMF was masked to group assignment. Participants in the control group received no further contact from the online Alive-PD system except reminders to complete a 3-month and 6-month online follow-up questionnaire. Because participants had consented only to a 6-month delay before they could start the intervention, only the 3-month and 6-month results constitute the randomized trial portion of the study.

The Alive-PD Intervention

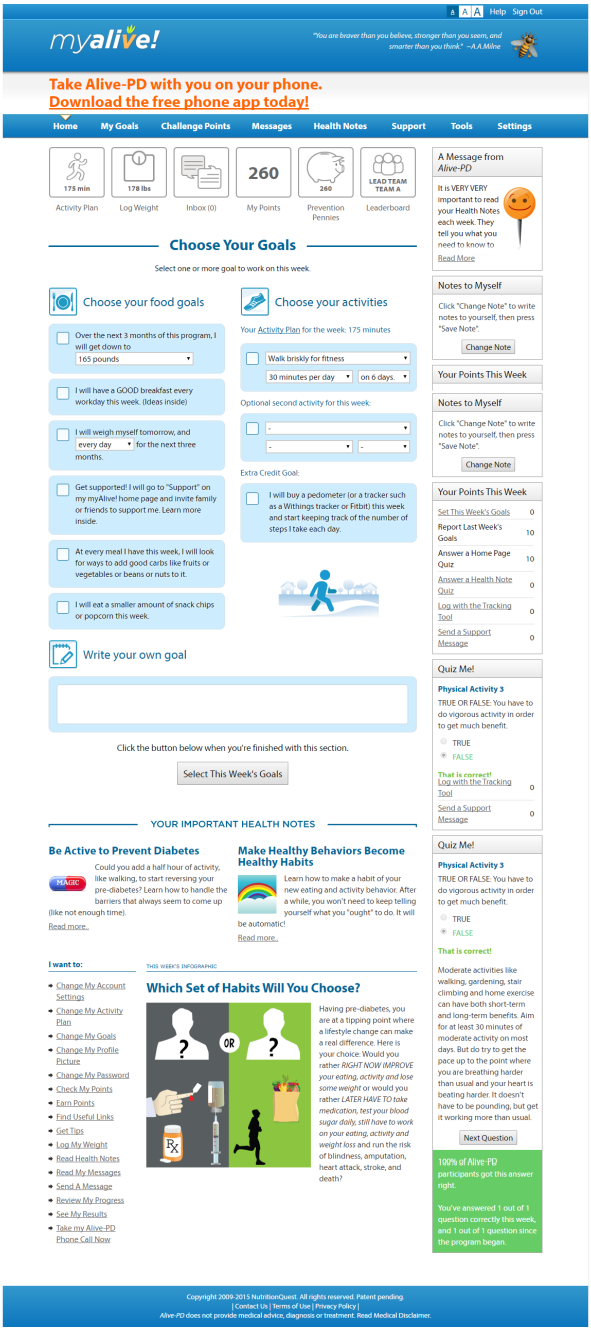
The program has been described in detail elsewhere [19]. Briefly, Alive-PD provides a 1-year program of regular contact and goal setting, weekly in the first 6 months and biweekly thereafter, plus midweek automated email and mobile phone reminders. The program includes individually tailored weekly goal setting and other activities delivered via Web and email supplemented by automated interactive voice response (IVR) phone calls and a supportive mobile phone app. Alive-PD was developed with input from, and was reviewed by, diabetes educators, endocrinologists, registered dietitians, and psychological experts in health behavior change. All features and contacts are completely automated and algorithm-driven, with no personal contact or coaching either in-person or remotely. See Figure 1 and Multimedia Appendix 2 for screenshots and other information.

The goal of the Alive-PD program is to improve glycemic control and reduce diabetes risk through lasting changes in physical activity and eating habits. Weight loss is encouraged and tracked as one of the changes that can reduce diabetes risk, although it is not the primary emphasis. For physical activity, participants set long-term goals of 150 to 300 minutes of aerobic activity per week depending on reported levels at baseline and on subsequent program participation. Resistance training is encouraged as well. For eating behaviors, the focus is on decreasing added sugars and refined carbohydrates, decreasing saturated and trans fats, and increasing fruit and vegetables. Changes in food type and reduction in portion size is emphasized as a means of reducing energy intake rather than specific calorie targets or counting. Psychosocial issues important in behavior change are addressed, including managing stress and sleep, staying motivated, addressing negative thoughts, modifying one's environment to support desired changes, and other topics addressed in the DPP curriculum [21].

These objectives are achieved through a system of weekly individually tailored goal setting. Based on a detailed initial questionnaire on current dietary and activity habits, and on the participant's subsequent interactions, the program recommends multiple weekly personally relevant small-step goals. Participants work on both increased physical activity and improved dietary habits each week, as well as occasional psychosocial goals. In addition to weekly personally tailored goals, the system provides tools for tracking weight, eating, and

physical activity; weekly health information on diabetes and strategies for preventing it; quizzes; social support through virtual teams and a participant messaging system; feedback on reported diet and activity and on success or failure of goal achievement; weekly reminders; and other features. Engagement is promoted through a points system with modest monetary rewards and team competition. During the first 6 months, participants are reminded automatically if they have not chosen a goal for 2 weeks using data from the online system.

Figure 1. Screenshot of Alive-PD personal home page.



An email initiates the choice of weekly goals, provides a link to the participant's Web page, and is followed up by a midweek reminder. An Android and iPhone app also permits the participant to select weekly goals, report on progress, and set mobile phone reminders. Automated motivational phone coaching is provided biweekly through IVR technology with interactions tailored to each individual's participation status, barriers, and primary motivations.

These strategies and features are based on established principles derived from several bodies of behavior change research. The basic objective, derived from learning theory and other habit formation research [22-24], is to have participants gradually incorporate new eating and physical activity behaviors into their daily lives until these behaviors are both habitual and substantial enough to reduce diabetes risk. To accomplish that objective, a variety of strategies are employed to sustain involvement with the program itself and, more importantly, to sustain the gradual incorporation of new healthier behaviors. The strategies are consistent with several bodies of research, including models centering on cues and triggers [25,26], social cognitive theory

[27,28], the theory of planned behavior [29], behavioral economics [26,30], and positive psychology [31,32]. For a more detailed description of the program, refer to the published protocol and program description [19].

Subsequent Clinic Visits

Participants in the intervention and control groups returned for clinic visits at 3 and 6 months, at which time the laboratory and biometric measurements described previously were repeated by trained staff unaware of treatment assignment. Active monitoring of adverse events was achieved by asking participants about sickness or injury at each clinic visit. At the 6-month visit, additional funding made it possible to invite participants to continue the program for another 6 months, although the randomized trial segment ended at 6 months. Those in the intervention group continued in that arm. Those in the control group were transferred to the active Alive-PD intervention program per the original consent. Participants who consented to the extension were seen at additional clinic visits at 9 and 12 months.

Statistical Methods

Intention-to-treat (ITT) analyses of change in HbA_{1c}, fasting glucose, and weight were prespecified. Baseline characteristics were compared by chi-square tests for categorical variables and *t* tests for continuous variables. Mean between-group treatment differences in outcomes were evaluated by ITT analysis using linear regression approaches. In all models, change in the outcome of interest (eg, HbA_{1c}) was the dependent variable with treatment group the main predictor (independent) variable and baseline value of the outcome variable as a covariate. Missing values in the dependent variable were imputed using the approach of Heckman et al [33,34], in which variables need not be assumed to be missing at random (MAR). This approach corrects for the bias in estimates of change that may arise from participants failing to complete the follow-up clinic visits. We examined potential interactions with treatment group by variables that were expected a priori to be potential effect modifiers (sex, race/ethnicity, age, and BMI category) by inclusion of a cross-product term in the model. No significant interactions were found. Adjustment for age, sex, BMI, and race/ethnicity did not materially alter the results. Dichotomous outcomes (eg, achievement of 5% weight loss) were evaluated by chi-square tests after confirming the absence of interactions using logistic regression. For comparability with other studies, we also conducted subgroup analyses on participants who were prediabetic by HbA_{1c} at baseline.

Results

Participant Randomization and Retention

A total of 340 participants met study eligibility criteria and were randomized. One participant randomized to the intervention group developed a metabolic condition rendering glycemic markers uninterpretable and was excluded from analysis, leaving 339 randomized participants.

Study retention and participation in biometric assessment visits was high; 89.1% (302/339) completed the 3-month follow-up assessment and 86.1% (292/339) completed the 6-month follow-up assessment. Of the 47 study participants that did not complete the 6-month follow-up (20 control, 27 intervention), 9 were lost to follow-up and 38 withdrew from the study. Reported adverse events were minor and all were considered to be unrelated to study participation. There were no significant differences in adverse events between treatment groups at either the 3-month or the 6-month visit (data not shown). One participant in the control group was diagnosed with diabetes and withdrew from the study; this participant did not provide follow-up measurements, but was included in the ITT analysis. No participants were prescribed metformin or other diabetes medications during the study.

Baseline Characteristics

Participants were a mean age of 55 (SD 8.9, range 31-70) years with a mean BMI of 31.1 (SD 4.4) kg/m^2 (Table 1). The majority (68.7%, 233/339) were male. Mean fasting glucose was at the low end of the prediabetic range (mean 6.1, SD 0.5 mmol/L or mean 109.9, SD 8.4 mg/dL) and mean HbA_{1c} was in the normal range (mean 5.6%, SD 0.3 or mean 38, SD 3.2 mmol/mol]). Metabolic syndrome was present in 68.1% (231/339) of participants. The study cohort was well educated; 82.9% (281/339) had a college degree or higher. The Framingham 8-year diabetes risk was 16.6% at baseline in both groups. The intervention and control groups were well balanced on baseline characteristics, although there was some imbalance for race/ethnicity, but it did not reach statistical significance (*P*=.07). This imbalance was due largely to a difference in Hispanic ethnicity (8.0%, 14/176 vs 4.3%, 7/163; *P*=.04). Due to this imbalance, race/ethnicity was examined for confounding and effect modification in all models.

Table 1. Baseline demographics and clinical characteristics.

Variable	All N=339	Control n=176	Intervention n=163	*P* [a]
Age (years), mean (SD)	55.0 (8.9)	54.9 (9.1)	55.0 (8.8)	.88
Female, n (%)	106 (31.3)	54 (30.7)	52 (31.9)	.81
College or above, n (%)	281 (82.9)	144 (81.8)	137 (84.1)	.59
Race/ethnicity, n (%) [b]				.07
White	229 (67.6)	120 (68.2)	109 (66.9)	
Hispanic	21 (6.2)	14 (8.0)	7 (4.3)	
Asian	70 (20.6)	29 (16.5)	41 (25.2)	
Other	19 (5.6)	13 (7.4)	6 (3.7)	
Metabolic syndrome, n (%)	231 (68.1)	121 (68.8)	110 (67.5)	.80
Weight (kg), mean (SD)	92.9 (15.8)	93.3 (16.6)	93.7 (14.9)	.68
BMI (kg/m^2), mean (SD)	31.2 (4.4)	31.2 (4.3)	31.1 (4.5)	.73
Waist circumference (cm), mean (SD)	102.8 (10.8)	103.1 (11.2)	102.5 (10.4)	.62
Glucose (mmol/L), mean (SD)	6.10 (0.5)	6.08 (0.5)	6.11 (0.5)	.57
Glucose (mg/dL), mean (SD)	109.9 (8.4)	109.6 (8.3)	110.1 (8.6)	.57
HbA_{1c}(%), mean (SD)	5.6 (0.3)	5.6 (0.3)	5.6 (0.3)	.90
HbA_{1c} (mmol/mol), mean (SD)	38.2 (3.2)	38.2 (3.1)	38.1 (3.3)	.90
Total cholesterol (mmol/L), mean (SD)	5.0 (0.8)	5.0 (0.9)	4.9 (0.8)	.82
LDL cholesterol (mmol/L), mean (SD)	3.0 (0.7)	3.0 (0.7)	3.0 (0.7)	.73
HDL cholesterol (mmol/L), mean (SD)	1.2 (0.4)	1.2 (0.3)	1.2 (0.4)	.34
Triglycerides (mmol/L), mean (SD)	1.6 (0.8)	1.7 (0.8)	1.6 (0.9)	.54
TG/HDL ratio, mean (SD)	3.5 (2.5)	3.6 (2.5)	3.4 (2.5)	.41
Blood pressure (mm Hg), mean (SD)				
Systolic	130.4 (14.7)	130.4 (14.5)	130.5 (15.0)	.95
Diastolic	82.3 (8.4)	82.6 (8.7)	82.0 (8.1)	.51
Framingham 8-year diabetes risk (%), mean (SD)	16.63 (10.67)	16.64 (10.78)	16.63 (10.58)	.99

[a] Significance of difference between intervention and control.

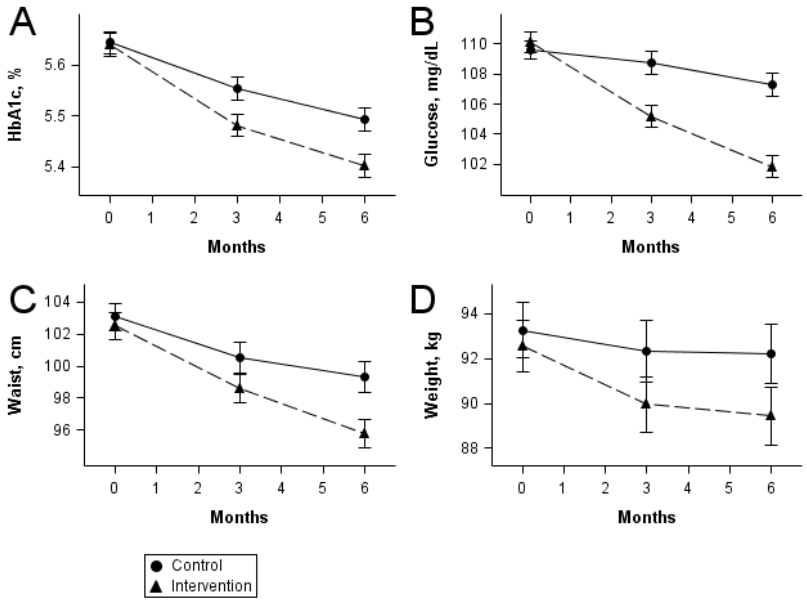
[b] Race and ethnicity as reported on online questionnaire. Native American/Alaskan, Native Hawaiian/Pacific Islander, more than one race, or "not reported" reported as "other."

Primary Outcomes

Significant decreases in HbA_{1c} and fasting glucose were observed in the intervention group by 3 months from baseline and declined further at 6 months (Figure 2).

In ITT analyses, which included all 339 participants, mean reductions in fasting glucose at 6 months from baseline were significantly greater in the intervention group (mean –0.41 mmol/L, 95% CI –0.44 to –0.12) than in the control group (mean –0.21 mmol/L, 95% CI –0.15 to –0.10; *P*<.001) (Table 2). Mean reductions in HbA_{1c} were also significantly greater in the intervention versus the control group (mean –0.26%, 95% CI –0.27 to –0.24 vs mean –0.18%, 95% CI –0.19 to –0.16; *P*<.001). No effect modification by race/ethnicity, age, sex, or BMI category was observed.

Figure 2. Changes in primary and secondary endpoints over time. Solid line: control; dashed line: intervention; error bars: ± standard error. A: Change in HbA1c. B: Change in fasting glucose. C: Change in waist circumference. D: Change in weight. At 6 months, all measures were significantly different between control and intervention groups (*P*<.001).



Although all participants had prediabetes at baseline by either HbA_{1c} or fasting glucose, only 44.8% (152/339) had prediabetes based on HbA_{1c}. In a subgroup analysis among those with prediabetes at baseline by HbA_{1c} (Table 2), the mean reduction in HbA_{1c} was greater than in the intervention group as a whole (mean –0.32%, 95% CI –0.38 to –0.26) and was significantly greater relative to the control group (mean –0.20%, 95% CI –0.25 to –0.15; *P*=.002).

Table 2. Change in clinical outcomes by treatment group.

Variable	Intention-to-treat,[a] change (95% CI)[b]			Prediabetic by HbA_{1c},[c] change (95% CI)[b]		
	Alive-PD n=163	Control n=176	*P*	Alive-PD n=60	Control n=69	*P*
Fasting glucose (mg/dL)	–7.36 (–7.85, –6.87)	–2.19 (–2.64, –1.73)	<.001	–7.38 (–9.40, –5.36)	–1.23 (–3.12, 0.65)	<.001
Fasting glucose (mmol/L)	–0.41 (–0.44, –0.38)	–0.12 (–0.15, –0.10)	<.001	–0.41 (–0.52, –0.30)	–0.07 (–0.17, 0.04)	<.001
HbA_{1c} (%)	–0.26 (–0.27, –0.24)	–0.18 (–0.19, –0.16)	<.001	–0.32 (–0.38, –0.27)	–0.20 (–0.25, –0.15)	.002
HbA_{1c} (mmol/mol)	–2.81 (–2.95, –2.66)	–1.93 (–2.06, –1.79)	<.001	–3.50 (–4.10, –2.90)	–2.15 (–2.71, –1.59)	.002
Weight (kg)	–3.26 (–3.26, –3.25)	–1.26 (–1.27, –1.26)	<.001	–3.56 (–4.42, –2.70)	–0.48 (–1.28, 0.32)	<.001
Weight loss (%)	–3.60 (–3.63, –3.57)	–1.32 (–1.36, –1.28)	<.001	–4.00 (–4.94, –3.07)	–0.53 (–1.40, 0.34)	<.001
BMI (kg/m^2)	–1.05 (–1.06, –1.05)	–0.39 (–0.39, –0.38)	<.001	–1.19 (–1.47, –0.90)	–0.17 (–0.43, 0.10)	<.001
Waist (cm)	–4.56 (–4.69, –4.43)	–2.22 (–2.36, –2.09)	<.001	–7.23 (–8.99, –5.47)	–2.73 (–4.37, –1.10)	<.001
TG/HDL ratio	–0.21 (–0.30, –0.12)	0.21 (0.12,0.29)	.04	–0.43 (–0.85, –0.02)	0.12 (–0.27, 0.51)	.06

[a] Imputation of missing dependent variables using Heckman/QLIM.

[b] 95% confidence limits from least squares means from models of following form: change=baseline + treatment group.

[c] Participants prediabetic by HbA_{1c} at baseline and providing complete data.

Secondary Outcomes

In the ITT analysis, reduction in weight, BMI, waist circumference, and TG/HDL ratio were all significantly greater in the intervention group than the control group (Table 2). The intervention group lost a mean 3.26 kg (95% CI –3.26 to –3.25) compared to 1.26 kg (95% CI –1.27 to –1.26) in the control group (*P*<.001). Mean BMI was reduced by 1.05 kg/m^2 (95% CI –1.06 to –1.05) and 0.39 kg/m^2 (95% CI –0.39 to –0.38) in the intervention and control groups, respectively (*P*<.001). The mean reduction in waist circumference in the intervention group was 4.56 cm (95% CI –4.69 to –4.43) compared to 2.22 cm (95% CI –2.36 to –2.09) in the control group (*P*<.001). In

addition, the ratio of TG/HDL was significantly reduced in the intervention group in contrast to the increase seen in the control group (mean –0.21, 95% CI –0.30 to –0.12 vs mean 0.21, 95% CI 0.12-0.29; *P*=.04).

The proportion of participants, by treatment group, meeting specific thresholds are shown in Figure 3. At 6 months, 35.3% (48/136) of the intervention group had achieved at least a 5% weight loss compared to 8.3% (13/156) of controls (Figure 3A). Among those who were prediabetic by fasting glucose at baseline, 40.5% (49/121) of intervention participants had achieved a normal fasting glucose compared to 17.7% (26/147) of controls (Figure 3B). Among participants who had metabolic syndrome at baseline, 46.5% (40/86) of those in the intervention group no longer had metabolic syndrome at 6 months compared with 20.0% (22/110) of controls (Figure 3C). BMI was reduced by at least 1 kg/m^2 in 44.9% (61/136) of intervention participants compared with 18.6% (29/156) of control participants (Figure 3D). All these differences between the intervention and control groups were significant at *P*<.001.

There was a significantly greater reduction in Framingham 8-year diabetes risk in the intervention versus the control group (*P*<.001) in the ITT sample (Figure 4). In both groups, the baseline diabetes risk was 16%. At 6 months, it was 11.00% (95% CI 10.08-11.92) in the intervention group and 14.59% (95% CI 13.64-15.54) in the control group.

Figure 3. Proportion achieving secondary endpoint thresholds at 6 months. Error bars not shown because all differences between control and intervention were *P*<.001. A: Percentage with ≥5% weight loss (complete data: n=156 control, n=136 intervention). B: Percentage who moved to normal fasting glucose (from ≥100 mg/dL to <100 mg/dL) (denominator: n=150 control, n=126 intervention). C: Percentage who moved from having metabolic syndrome to not having metabolic syndrome (denominator: n=110 control, n=86 intervention). D: Percentage whose BMI decreased by 1 kg/m^2 (denominator: n=156 control, n=136 intervention).

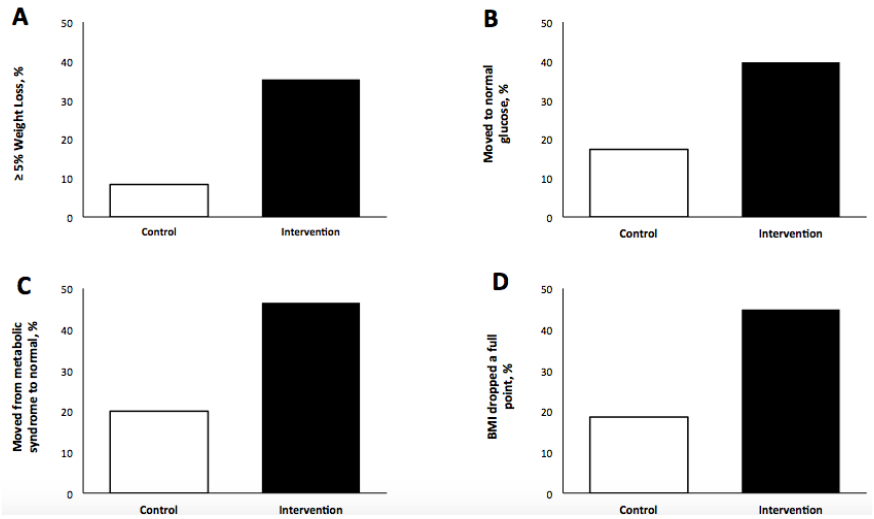
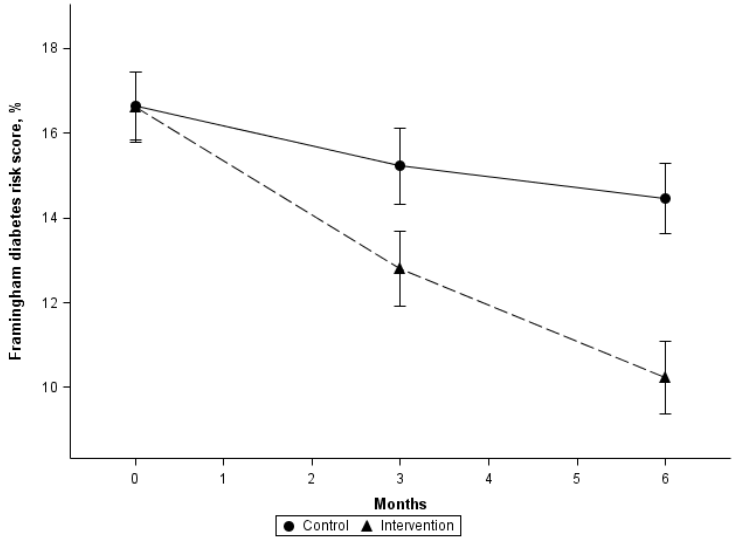


Figure 4. Change in Framingham 8-year diabetes risk.



Case Report on Participants in the Diabetic Range by Fasting Glucose

Alive-PD was designed to assist persons with prediabetes. However, lifestyle behavior change is also an essential intervention for persons who are newly diagnosed with diabetes. Thus, information about results in the 8 participants in our sample who had a fasting glucose in the diabetes range at baseline is of interest (they were all cleared by their physicians for participation in the study). Of the 5 in the intervention group, one had a decrease in fasting glucose to the normal range (<100 mg/dL) and the other 4 had a decrease in fasting glucose to the prediabetic range (<126 mg/dL) after the 6-month intervention period. None of the 3 participants in the control group had decreases in glucose outside of the diabetic range.

Process Measures and Other Behaviors

We assessed program participation by evaluating the points each participant earned through interacting with the program components each week and by assessing the participants' weekly goal setting behaviors. Participation in the online Alive-PD program features was high. Intervention participants (ITT population, n=163) set behavioral goals or otherwise interacted with the online Alive-PD program in a median of 17 (IQR 14) of the 24 weeks (70.8% of the weeks). In all, 87.1% (142/163) interacted with the program in 4 or more of the 24 weeks and 70.6% (115/163) were still interacting with the program in the last month of the 6-month period. Participants accomplished a median of 35 goals (IQR 107) in the 24-week period or approximately 1.5 goals per week. Intervention participants reported that they spent approximately 15 minutes interacting with the program in a typical week.

The intervention group experienced significant improvements in self-reported physical activity, dietary habits, sleep, fatigue, and self-confidence relative to the control group (P<.001) (data not shown). A more detailed analysis of changes in physical activity, diet, self-confidence, and other psychosocial factors will be reported elsewhere.

Discussion

In this randomized controlled trial, the fully automated Alive-PD program was effective in improving glycemic control and body weight, and in reducing 8-year diabetes risk. In ITT analyses, the intervention group achieved reductions in fasting glucose of –41 mmol/L (–7.36 mg/dL) and in HbA_{1c} of –0.26% (–3 mmol/mol), both statistically significantly superior to changes in the control group. In addition, intervention group participants lost a mean 3.26 kg over 6 months, in ITT analyses, and 35% of the intervention group lost 5% or more of initial body weight, both significantly superior to the control group.

Previous Research on Weight Loss in Diabetes Prevention or Weight Loss Programs

Numerous reviews of weight loss or translational diabetes prevention programs have been conducted [9,12,35-42] covering a range of delivery methods. Interventions using in-person or group approaches have achieved average weight losses of approximately 3% to 4% in reviews and meta-analyses [9,37], although some individual studies have reached weight losses of more than 6% in the intervention group [6,43].

For wider reach, however, many interventions have combined coaches with some form of technology. In a 2015 review of 16 studies of technology-assisted programs for weight loss in primary care, Levine et al [12] found a median weight loss of -2.7 kg in intervention groups of 12 programs that combined human with technological methods. Ali et al [9] found a mean loss of 4.2% of body weight in electronic media-assisted programs.

Interventions delivered entirely by electronic media, primarily for weight loss, have also been reviewed. Hartmann-Boyce et al [13] conducted a meta-analysis of 23 randomized trials of "self-help interventions" for weight loss in overweight or obese adults. Programs were not eligible for inclusion if they used any form of person-to-person assistance by counselors or health professionals. The analysis found a mean difference between intervention and comparison groups of -1.85 kg (95% CI -2.86 to -0.83) at 6 months. Three programs using eHealth technologies that were not included in the Hartmann-Boyce review were found by Hutchesson et al [14] to have a mean difference of -1.5 kg. Three other fully automated studies from the Levine review [12] found a mean weight loss in the intervention group of 2.5 kg. One recent trial not included in previous reviews [44] was fully automated with the exception of a 60-minute baseline visit at which participants were given weight loss, calorie and physical activity goals, and taught behavioral skills. A weight loss of 5.4 kg was observed at 6 months.

The effect of Internet-based interventions on change in waist circumference has also been examined in a meta-analysis. Seo and Niu [45] found a mean change of -2.99 cm (95% CI −3.68 to −2.30).

Previous Research on the Effect of Fully Automated Programs on Glycemic Markers

With few exceptions, most studies of diabetes prevention or weight loss interventions using fully automated programs have not measured or reported on changes in glycemic markers. One review found "minimal" changes in glycemic markers across the reviewed studies, with a median change in fasting glucose of -0.2 mmol/L [37] and another found a mean change of -0.1 mmol/L [38]. For HbA_{1c}, Dunkley et al [38] found pooled changes of -0.13% and Johnson et al [37] found a median change of -0.05%.

The treatment effects for Turnaround Health's Alive-PD program are consistent with and, in most cases, somewhat larger than the results summarized in the preceding meta-analyses. This is true for weight loss (–3.26 kg), percent weight loss (–3.60%), waist circumference (–4.56 cm), and the glycemic markers HbA_{1c} (–0.26%) and fasting glucose (–0.41 mmol/L), all in ITT analyses.

The Diabetes Prevention Recognition Program

The Centers for Disease Control and Prevention (CDC) Diabetes Prevention Recognition Program (DPRP) is intended to recognize organizations that have demonstrated their ability to

deliver a proven type 2 diabetes prevention lifestyle intervention [46]. The CDC recently updated the requirements for recognition to include programs delivered "virtually" provided they meet other criteria. Turnaround Health's Alive-PD program is listed on the CDC website [47]. As of August 1, 2015, it is the only such program with evidence of effectiveness from a randomized controlled trial and the only study with ITT analysis.

Features Promoting Effectiveness

A number of authors have attempted to identify or summarize what features of a behavioral intervention may be associated with its effectiveness [40,42,48]. The following have all been identified as contributors to effectiveness: goal setting, self-monitoring, tailoring and tailored feedback, reminders, social support, and a structured program employing behavior change principles. Khaylis et al [48] also listed feedback by a counselor as an important feature, but noted that computer-automated email feedback has been as effective as human email counseling in at least one study. With the exception of human counseling, all these features are incorporated into the Alive-PD program. In addition, Alive-PD added some gamification features, such as a points system, team competition, and monetary rewards, to enhance engagement and retention.

Research is underway to explore which features of Alive-PD may be more important or beneficial. Although all participants were exposed to all these components (goal setting, messaging, etc), different participants engaged in them to different extents. For example, 38.7% (63/163) never logged their weight or physical activity, whereas 12.3% (20/163) logged their weight or activity in 21 or more of the 24 weeks. Mediation analyses are underway. However, it is worth noting that participants are individuals with varying interests and motivations. Some people appreciate being on a team whereas others dislike it and the same can be said of other components. Alive-PD was intentionally designed to provide an array of components to engage the widest range of different interests, learning styles, and available time.

In addition to the potential role of features of an intervention, it is also of considerable interest to explore what behaviors and specific changes contributed to the study outcomes. Recent literature has discussed the relative roles of types of macronutrients (fats vs carbohydrates), physical activity, and weight loss [49-51]. The Alive-PD program promoted, and achieved, increases in physical activity, reductions in refined carbohydrates, reductions in saturated and trans fats, and increases in fruits and vegetables. Changes in specific foods were also encouraged, such as nuts, legumes, and olive oil. Participants in the intervention group undertook these changes to varying degrees. In future analyses, we will examine the effect of these variations on changes in glycemic markers and weight. For example, there was a significant reduction in HbA_{1c}, even among those who did not achieve 5% weight loss. We plan to explore factors that contributed to glycemic improvements in the absence of major weight loss.

Limitations

The fully automated nature of the Alive-PD program is both a strength and a limitation. Some people need and respond better to human interaction and support, and effect sizes might be greater if combined with human support. In addition, because the intervention is delivered by email, Internet, and mobile phone, it may have limited reach for those who do not have Internet access or who are not technologically proficient. Although its reach is somewhat limited in that respect, 87% of American adults used the Internet as of 2014, including more than 80% of African Americans and Hispanics [52]. These technologies are nearly ubiquitous in society and allow for convenient program access at home or through mobile devices. At the same time, the fully automated characteristic of the program is beneficial for several reasons. There is a guarantee of 100% fidelity to the design and content in future administrations and enhancements can be readily incorporated. Because it is fully automated, this commercial program can be delivered at low cost and with wide reach. Additionally, organizations using it would require no or minimal staff.

Although the Alive-PD program provides a 1-year intervention, the randomized trial analysis was for only a 6-month period. This was due to initial funding limitations and the desire to enhance enrollment of these persons at high risk of developing diabetes by assuring them that they would be given access to the active program in a reasonable period of time. It will be important to follow study participants for a full year to determine whether the trends seen in Figure 2 are maintained.

Study participants were relatively well educated and two-thirds were non-Hispanic white. Thus, the generalizability to less educated individuals and those of race/ethnic minority groups remains to be investigated. However, it is notable that the subgroup with postgraduate or professional degrees achieved less improvement in glycemic markers than those with lower educational levels (data not shown.) The sample did include a substantial number of Asians (21% of the study cohort) including South Asians, a group for which type 2 diabetes is especially common. Although analyses indicated no significant differences in treatment effects by ethnicity, more research is needed to confirm effectiveness in minority groups.

Clinical Relevance

The decrease in fasting glucose in the intervention group (–7.36 mg/dL or –0.41 mmol/L) was clinically meaningful and substantial. The decrease in HbA_{1c} was modest (–0.26% or –2.81 mmol/mol in the ITT analysis and –0.32% or –3.5 mmol/mol in those prediabetic by HbA_{1c}), but significantly greater than in controls. We note that baseline levels of HbA_{1c} were low in the study cohort. Indeed, mean HbA_{1c} at baseline was in the normal range and only 45% were prediabetic by the HbA_{1c} definition. As a result, the magnitude of the average treatment effect was not as large as might be expected in patients with higher values in the prediabetic range or in those with diabetes. Weight loss was 4% of baseline weight among those prediabetic by HbA_{1c} (Table 2) and increased with increasing participation in the program and higher baseline weight (data not shown). As noted, the primary objective was to lower glycemic markers, a direct measure of reduced diabetes risk, and this appears to have been achieved despite the relatively modest weight loss. The Alive-PD group's decreases in HbA_{1c}

and fasting glucose were greater than those seen in the Diabetes Prevention Program [5] Lifestyle group at 6 months (HbA_{1c}: -0.26% vs -0.09%; fasting glucose: -7.36 mg/dL vs 4.59 mg/dL, respectively), despite the fact that the Alive-PD group's weight loss was not as great as that seen in DPP (3.26 kg vs -6.5 kg) [5].

More than two-thirds of enrollees were male, a different sex distribution than is usually seen in health interventions (the DPP had 68% female participants) [5]. The electronic format may have had more appeal for men than a series of group or personal interactions. There was not a significant interaction between sex and treatment effect, and treatment effects were not significantly different by sex for HbA_{1c}, fasting glucose, or weight.

Summary

In summary, Alive-PD was effective in improving markers of glycemic control and body weight in patients with prediabetes. As noted by Cefalu et al [4], the driving force behind the increased economic burden of diabetes is increased prevalence. Therefore, engaging as many as possible of the nation's 86 million adults with prediabetes with a variety of cost-effective interventions is an urgent priority. Effective fully automated technologies such as Alive-PD represent one of those strategies, with the potential of serving large numbers of persons at risk of progression to diabetes.

Acknowledgments

Research reported in this publication was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Number R44NR012617. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

GB designed and led the study. KA was principal investigator of the Palo Alto Medical Foundation Research Institute subaward for the clinical component and RJR was coinvestigator of the subaward. GB, KA, and RJR cowrote the manuscript and TJB, CHB, LP, and MD contributed to the writing of the manuscript. TJB, GB, CHB, and HAC designed and implemented the Alive-PD program. DH conducted and led the systems engineering. GB and RJR conducted the analyses and MLH reviewed the analyses. MD researched data and contributed to program implementation.

Conflicts of Interest

GB, CHB, and TJB are the owners of Turnaround Health and NutritionQuest, the developers of Alive-PD. KA, RJR, LP, MD, DH, HAC, and MLH have no conflicts of interest.

Multimedia Appendix 1

CONSORT-EHEALTH checklist.

[PDF File (Adobe PDF File), 2MB - jmir_v17i10e240_app1.pdf]

Multimedia Appendix 2

Powerpoint presentation with screenshots of Alive-PD.

[PPT File (Microsoft PowerPoint Presentation), 4MB - jmir_v17i10e240_app2.ppt]

References

1. American Diabetes Association. 2014 Jun 10. Statistics About Diabetes URL: http://www.diabetes.org/diabetes-basics/statistics/ [accessed 2015-06-29] [WebCite Cache ID 6ZeYrQp5D]
2. Tabák AG, Herder C, Rathmann W, Brunner E, Kivimäki M. Prediabetes: a high-risk state for diabetes development. Lancet 2012 Jun 16;379(9833):2279-2290 [FREE Full text] [doi: 10.1016/S0140-6736(12)60283-9] [Medline: 22683128]
3. Dall TM, Yang W, Halder P, Pang B, Massoudi M, Wintfeld N, et al. The economic burden of elevated blood glucose levels in 2012: diagnosed and undiagnosed diabetes, gestational diabetes mellitus, and prediabetes. Diabetes Care 2014 Dec;37(12):3172-3179. [doi: 10.2337/dc14-1036] [Medline: 25414388]
4. Cefalu W, Petersen M, Ratner R. The alarming and rising costs of diabetes and prediabetes: a call for action!. Diabetes Care 2014 Dec;37(12):3137-3138. [doi: 10.2337/dc14-2329] [Medline: 25414386]
5. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 7;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]

6. Ackermann RT, Finch EA, Brizendine E, Zhou H, Marrero DG. Translating the Diabetes Prevention Program into the community. The DEPLOY Pilot Study. Am J Prev Med 2008 Oct;35(4):357-363 [FREE Full text] [doi: 10.1016/j.amepre.2008.06.035] [Medline: 18779029]
7. Whittemore R, Melkus G, Wagner J, Dziura J, Northrup V, Grey M. Translating the diabetes prevention program to primary care: a pilot study. Nurs Res 2009;58(1):2-12 [FREE Full text] [doi: 10.1097/NNR.0b013e31818fcef3] [Medline: 19092550]
8. Davis-Smith Y, Davis-Smith M, Boltri JM, Seale JP, Shellenberger S, Blalock T, et al. Implementing a diabetes prevention program in a rural African-American church. J Natl Med Assoc 2007 Apr;99(4):440-446. [Medline: 17444435]
9. Ali M, Echouffo-Tcheugui J, Williamson D. How effective were lifestyle interventions in real-world settings that were modeled on the Diabetes Prevention Program? Health Aff (Millwood) 2012 Jan;31(1):67-75 [FREE Full text] [doi: 10.1377/hlthaff.2011.1009] [Medline: 22232096]
10. Lawlor M, Blackwell C, Isom S, Katula JA, Vitolins MZ, Morgan TM, et al. Cost of a group translation of the Diabetes Prevention Program: Healthy Living Partnerships to Prevent Diabetes. Am J Prev Med 2013 Apr;44(4 Suppl 4):S381-S389 [FREE Full text] [doi: 10.1016/j.amepre.2012.12.016] [Medline: 23498303]
11. Vojta D, Koehler TB, Longjohn M, Lever JA, Caputo NF. A coordinated national model for diabetes prevention: linking health systems to an evidence-based community program. Am J Prev Med 2013 Apr;44(4 Suppl 4):S301-S306. [doi: 10.1016/j.amepre.2012.12.018] [Medline: 23498291]
12. Levine DM, Savarimuthu S, Squires A, Nicholson J, Jay M. Technology-assisted weight loss interventions in primary care: a systematic review. J Gen Intern Med 2015 Jan;30(1):107-117. [doi: 10.1007/s11606-014-2987-6] [Medline: 25134692]
13. Hartmann-Boyce J, Jebb S, Fletcher B, Aveyard P. Self-help for weight loss in overweight and obese adults: systematic review and meta-analysis. Am J Public Health 2015 Mar;105(3):e43-e57. [doi: 10.2105/AJPH.2014.302389] [Medline: 25602873]
14. Hutchesson M, Rollo M, 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]
15. McLaughlin T, Reaven G, Abbasi F, Lamendola C, Saad M, Waters D, et al. Is there a simple way to identify insulin-resistant individuals at increased risk of cardiovascular disease? Am J Cardiol 2005 Aug 1;96(3):399-404. [doi: 10.1016/j.amjcard.2005.03.085] [Medline: 16054467]
16. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, International Diabetes Federation Task Force on EpidemiologyPrevention, National Heart, Lung, and Blood Institute, International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009 Oct 20;120(16):1640-1645 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192644] [Medline: 19805654]
17. Framingham Heart Study. 2015. Diabetes Risk Score URL: http://www.framinghamheartstudy.org/risk-functions/diabetes/index.php [accessed 2015-10-18] [WebCite Cache ID 6cNad0tXF]
18. Davis NJ, Tomuta N, Schechter C, Isasi CR, Segal-Isaacson CJ, Stein D, et al. Comparative study of the effects of a 1-year dietary intervention of a low-carbohydrate diet versus a low-fat diet on weight and glycemic control in type 2 diabetes. Diabetes Care 2009 Jul;32(7):1147-1152 [FREE Full text] [doi: 10.2337/dc08-2108] [Medline: 19366978]
19. Block G, Azar KM, Block TJ, Romanelli RJ, Carpenter H, Hopkins D, et al. A fully automated diabetes prevention program, Alive-PD: program design and randomized controlled trial protocol. JMIR Res Protoc 2015;4(1):e3 [FREE Full text] [doi: 10.2196/resprot.4046] [Medline: 25608692]
20. Hsu W, Araneta MR, Kanaya A, Chiang J, Fujimoto W. BMI cut points to identify at-risk Asian Americans for type 2 diabetes screening. Diabetes Care 2015 Jan;38(1):150-158. [doi: 10.2337/dc14-2391] [Medline: 25538311]
21. Centers for Disease Control and Prevention. 2015. The CDC Diabetes Prevention Program (CDC DPP) Curriculum URL: http://www.cdc.gov/diabetes/prevention/recognition/curriculum.htm [accessed 2015-06-29] [WebCite Cache ID 6ZeZBvypB]
22. Duhigg C. Power of Habit: Why We Do What We Do, and How to Change. London: Heinemann Educational Books; 2012.
23. Hull CL. Essentials of Behavior. New Haven: Yale University Press; 1951.
24. Dickinson A. Actions and habits: the development of behavioural autonomy. Phil Trans R Soc Lond B 1985 Feb 13;308(1135):67-78. [doi: 10.1098/rstb.1985.0010]
25. Fogg GJ. Stanford Persuasive Tech Lab. 2014. URL: https://captology.stanford.edu/ [accessed 2015-06-29] [WebCite Cache ID 6ZeZ3Rnou]
26. Wansink B. Mindless Eating: Why We Eat More Than We Think. New York: Bantam; 2006.
27. Brownell K, Marlatt G, Lichtenstein E, Wilson G. Understanding and preventing relapse. Am Psychol 1986 Jul;41(7):765-782. [Medline: 3527003]
28. Beck J. The Complete Beck Diet for Life: The 5-Stage Program for Permanent Weight Loss. Birmingham, AL: Oxmoor House; 2008.
29. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]

30. Heshmat S. Eating Behavior and Obesity: Behavioral Economics Strategies for Health Professionals. New York: Springer Publishing Company; 2011.
31. Cornum R, Matthews M, Seligman ME. Comprehensive soldier fitness: building resilience in a challenging institutional context. Am Psychol 2011 Jan;66(1):4-9. [doi: 10.1037/a0021420] [Medline: 21219042]
32. Seligman ME, Railton P, Baumeister RF, Sripada C. Navigating into the future or driven by the past. Perspect Psychol Sci 2013 Mar;8(2):119-141. [doi: 10.1177/1745691612474317] [Medline: 26172493]
33. Heckman J. The common structure of statistical models of truncated, sample selection and limited dependent variables, and a simple estimator of such models. In: Annals of Economic and Social Measurement, Volume 5, number 4. Cambridge, MA: The National Bureau of Economic Research; 1976:475-492.
34. Allison P. Paper 312-. 2012. Handling Missing Data by Maximum Likelihood URL: http://www.statisticalhorizons.com/wp-content/uploads/MissingDataByML.pdf [accessed 2015-10-18] [WebCite Cache ID 6cNh0D2Wr]
35. Whittemore R. A systematic review of the translational research on the Diabetes Prevention Program. Transl Behav Med 2011 Sep;1(3):480-491 [FREE Full text] [doi: 10.1007/s13142-011-0062-y] [Medline: 24073067]
36. Norris SL, Zhang X, Avenell A, Gregg E, Bowman B, Schmid CH, et al. Long-term effectiveness of weight-loss interventions in adults with pre-diabetes: a review. Am J Prev Med 2005 Jan;28(1):126-139. [doi: 10.1016/j.amepre.2004.08.006] [Medline: 15626569]
37. Johnson M, Jones R, Freeman C, Woods HB, Gillett M, Goyder E, et al. Can diabetes prevention programmes be translated effectively into real-world settings and still deliver improved outcomes? A synthesis of evidence. Diabet Med 2013 Jan;30(1):3-15 [FREE Full text] [doi: 10.1111/dme.12018] [Medline: 22998334]
38. Dunkley AJ, Bodicoat DH, Greaves CJ, Russell C, Yates T, Davies MJ, et al. Diabetes prevention in the real world: effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes and of the impact of adherence to guideline recommendations: a systematic review and meta-analysis. Diabetes Care 2014 Apr;37(4):922-933. [doi: 10.2337/dc13-2195] [Medline: 24652723]
39. McTigue KM, Conroy MB. Use of the internet in the treatment of obesity and prevention of type 2 diabetes in primary care. Proc Nutr Soc 2013 Feb;72(1):98-108. [doi: 10.1017/S0029665112002777] [Medline: 23098133]
40. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
41. Allen J, Stephens J, Patel A. Technology-assisted weight management interventions: systematic review of clinical trials. Telemed J E Health 2014 Dec;20(12):1103-1120. [doi: 10.1089/tmj.2014.0030] [Medline: 25409001]
42. Tang J, Abraham C, Greaves C, Yates T. Self-directed interventions to promote weight loss: a systematic review of reviews. J Med Internet Res 2014;16(2):e58 [FREE Full text] [doi: 10.2196/jmir.2857] [Medline: 24554464]
43. Ma J, Yank V, Xiao L, Lavori PW, Wilson SR, Rosas LG, et al. Translating the Diabetes Prevention Program lifestyle intervention for weight loss into primary care: a randomized trial. JAMA Intern Med 2013 Jan 28;173(2):113-121 [FREE Full text] [doi: 10.1001/2013.jamainternmed.987] [Medline: 23229846]
44. Thomas JG, Leahey TM, Wing RR. An automated internet behavioral weight-loss program by physician referral: a randomized controlled trial. Diabetes Care 2015 Jan;38(1):9-15. [doi: 10.2337/dc14-1474] [Medline: 25404659]
45. Seo D, Niu J. Evaluation of Internet-based interventions on waist circumference reduction: a meta-analysis. J Med Internet Res 2015;17(7):e181 [FREE Full text] [doi: 10.2196/jmir.3921] [Medline: 26199208]
46. Centers for Disease Control and Prevention. 2015. Diabetes Prevention Recognition Program Standards and Operating Procedures URL: http://www.cdc.gov/diabetes/prevention/recognition/standards.htm [accessed 2015-10-18] [WebCite Cache ID 6cNoRsg3z]
47. Centers for Disease Control and Prevention. 2015. Registry of Recognized Programs, Virtual, Online or Combination of In-person/Online Programs URL: https://nccd.cdc.gov/DDT_DPRP/Registry.aspx [accessed 2015-08-04] [WebCite Cache ID 6aXdCnzHn]
48. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. Telemed J E Health 2010 Nov;16(9):931-938 [FREE Full text] [doi: 10.1089/tmj.2010.0065] [Medline: 21091286]
49. Winett R, Davy B, Marinik E, Savla J, Winett SG, Phillips SM, et al. Developing a new treatment paradigm for disease prevention and healthy aging. Transl Behav Med 2014 Mar;4(1):117-123 [FREE Full text] [doi: 10.1007/s13142-013-0225-0] [Medline: 24653782]
50. Salas-Salvadó J, Martinez-González M, Bulló M, Ros E. The role of diet in the prevention of type 2 diabetes. Nutr Metab Cardiovasc Dis 2011 Sep;21 Suppl 2:B32-B48. [doi: 10.1016/j.numecd.2011.03.009] [Medline: 21745730]
51. Risérus U, Willett W, Hu F. Dietary fats and prevention of type 2 diabetes. Prog Lipid Res 2009 Jan;48(1):44-51 [FREE Full text] [doi: 10.1016/j.plipres.2008.10.002] [Medline: 19032965]
52. Fox S, Rainie L. Pew Research Center. 2014 Feb 27. The Web at 25 in the US URL: http://www.pewinternet.org/2014/02/27/the-web-at-25-in-the-u-s/ [accessed 2015-06-29] [WebCite Cache ID 6ZehHJWHV]

Abbreviations

BMI: body mass index
CDC: Centers for Disease Control and Prevention
DPP: Diabetes Prevention Program
DPRP: Diabetes Prevention Recognition Program
HDL: high-density lipoprotein
ITT: intention-to-treat
IVR: interactive voice response
PAMF: Palo Alto Medical Foundation
TG: triglyceride

Edited by G Eysenbach; submitted 07.07.15; peer-reviewed by I Pires, H Payne; comments to author 31.07.15; revised version received 15.08.15; accepted 21.09.15; published 23.10.15.

Please cite as:
Block G, Azar KMJ, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, Dolginsky MS, Hudes ML, Palaniappan LP, Block CH
Diabetes Prevention and Weight Loss with a Fully Automated Behavioral Intervention by Email, Web, and Mobile Phone: A Randomized Controlled Trial Among Persons with Prediabetes
J Med Internet Res 2015;17(10):e240
URL: http://www.jmir.org/2015/10/e240/
doi:10.2196/jmir.4897
PMID:26499966

Original Paper

Web-Based STAR E-Learning Course Increases Empathy and Understanding in Dementia Caregivers: Results from a Randomized Controlled Trial in the Netherlands and the United Kingdom

Bart Hattink[1], MSc; Franka Meiland[1], PhD; Henriëtte van der Roest[1], PhD; Peter Kevern[2], PhD; Francesca Abiuso[3], PhD; Johan Bengtsson[4], MSc; Angele Giuliano[5], MSc; Annalise Duca[5], BSc; Jennifer Sanders[2], MSc; Fern Basnett[2], MSc; Chris Nugent[6], PhD; Paul Kingston[7], PhD; Rose-Marie Dröes[1], PhD

[1]VU University medical center Amsterdam, Department of Psychiatry, Amsterdam, Netherlands

[2]Staffordshire University, Faculty of Health Sciences, Stafford, United Kingdom

[3]Medea, Firenze, Italy

[4]Internit, Lulea, Sweden

[5]Across Limits, Hamrun, Malta

[6]University of Ulster, Faculty of Computing and Engineering, Newtownabbey, United Kingdom

[7]University of Chester, Department of Health & Social Care, Chester, United Kingdom

Corresponding Author:
Bart Hattink, MSc
VU University medical center Amsterdam
Department of Psychiatry
GGZ inGeest Dienst Onderzoek en Innovatie - Overschiestraat
Postbus 74077
Amsterdam, 1070BB
Netherlands
Phone: 31 207885622
Fax: 31 207885622
Email: b.hattink@vumc.nl

Abstract

Background: The doubling of the number of people with dementia in the coming decades coupled with the rapid decline in the working population in our graying society is expected to result in a large decrease in the number of professionals available to provide care to people with dementia. As a result, care will be supplied increasingly by untrained informal caregivers and volunteers. To promote effective care and avoid overburdening of untrained and trained caregivers, they must become properly skilled. To this end, the European Skills Training and Reskilling (STAR) project, which comprised experts from the domains of education, technology, and dementia care from 6 countries (the Netherlands, Sweden, Italy, Malta, Romania, and the United Kingdom), worked together to create and evaluate a multilingual e-learning tool. The STAR training portal provides dementia care training both for informal and formal caregivers.

Objective: The objective of the current study was to evaluate the user friendliness, usefulness, and impact of STAR with informal caregivers, volunteers, and professional caregivers.

Methods: For 2 to 4 months, the experimental group had access to the STAR training portal, a Web-based portal consisting of 8 modules, 2 of which had a basic level and 6 additional modules at intermediate and advanced levels. The experimental group also had access to online peer and expert communities for support and information exchange. The control group received free access to STAR after the research had ended. The STAR training portal was evaluated in a randomized controlled trial among informal caregivers and volunteers in addition to professional caregivers (N=142) in the Netherlands and the United Kingdom. Assessments were performed with self-assessed, online, standardized questionnaires at baseline and after 2 to 4 months. Primary outcome measures were user friendliness, usefulness, and impact of STAR on knowledge, attitudes, and approaches of caregivers regarding dementia. Secondary outcome measures were empathy, quality of life, burden, and caregivers' sense of competence.

Results: STAR was rated positively by all user groups on both usefulness and user friendliness. Significant effects were found on a person-centered care approach and on the total score on positive attitudes to dementia; both the experimental and the control group increased in score. Regarding empathy, significant improvements were found in the STAR training group on distress, empathic concern, and taking the perspective of the person with dementia. In the experimental group, however, there was a significant reduction in self-reported sense of competence.

Conclusions: The STAR training portal is a useful and user-friendly e-learning method, which has demonstrated its ability to provide significant positive effects on caregiver attitudes and empathy.

(J Med Internet Res 2015;17(10):e241) doi:10.2196/jmir.4025

KEYWORDS

dementia; caregivers; distance-learning; empathy

Introduction

The European Union (EU) is set to face major demographic challenges in the coming decades. Two main drivers for this are the (expected) doubling of the number of people with dementia and a rapid relative decline in the working population. In the Netherlands, for example, this is expected to change from a ratio of 1:42 for people with dementia to working people in 2010 to 1:16 in 2050 [1]. As a result, the task of caring for people with dementia will be provided increasingly by relatives or friends, the so-called informal caregivers, who provide this care unpaid and generally with minimal or no professional assistance. Additionally, many EU countries draft their health care policies toward an increased use of volunteers in the provision of care in addition to prolonging community-based dementia care. Therefore, to sustain and promote effective care for people with dementia, to avoid overburdening of informal and professional caregivers, and to prevent premature admission of people with dementia to long-term care settings, caregivers need to become properly skilled and feel competent in their care provision.

In an attempt to address this, e-learning interventions could prove to be a useful tool to assist informal caregivers, untrained volunteers, and professionals by offering them relevant education, training, and support [2,3] at a significantly lower cost than through face-to-face training or print distribution [4]. Interventions offered through the Internet are likely to have a lower threshold for participation given that participants can use these interventions at any time they wish, from their own homes, and with little effort. This will also help to offer access to people who would otherwise be put off by long travel times, avoid costs for visiting regular teaching sessions (eg, people living in remote areas [5]), or to people who cannot leave their home due to their caregiving role. Finally, the possibilities of the Internet allow for effective use of multimedia delivery of information (eg, graphics, animations, and interactive course material), which has been reported to enhance learning and make the material more attractive during the process of engagement [6]. Recent research has found beneficial effects from Internet-based interventions. A Cochrane review in 2005 found that "interactive health communication applications" were effective for increasing knowledge and may improve outcomes in patients and caregivers [7]. A typical means for distributing interactive health communication apps is the Internet. In another review, it was found that personalized (tailored to the individual) Internet-based interventions led to improved health in users [8].

Pilot studies offering an Internet-based program of learning for dementia caregivers found that the caregivers who evaluated it reported it as useful, educational, and convenient [9] and found positive results relating to knowledge, attitudes, self-efficacy, and empathy [10]. Among professional caregivers, e-learning was also found to be enjoyable and was reported to help acquire new skills for collaboration among professionals [11]. E-learning was also found to help staff in nursing homes to gain specific skills, such as delirium screening [12]. A review of the state-of-the-art of online course provision for providing care for people with dementia in 2011 for 4 European countries (Netherlands, the United Kingdom, Malta, and Romania) showed that 14% of the dementia courses were offered online in the Netherlands, 17% in the United Kingdom, and in both Malta and Romania there were no online courses relating to care provision for persons with dementia [13]. These findings formed the basis for the development and evaluation of a multilingual online learning platform for all types of dementia caregivers, within the EU Skills Training and Reskilling (STAR) project [14].

The European STAR project (2010-2014), funded by the European Commission in the Leonardo da Vinci Life Long Learning Programme, aimed to improve the knowledge about dementia for informal caregivers, volunteers, and professionals in dementia care by developing and evaluating an online training program in different languages and at different difficulty levels. The course content was developed from 3 theoretical perspectives: (1) the medical model of dementia, including information on types of dementia, symptoms, and diagnostics based on the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) [15]; (2) the perspective of functional consequences in daily life based on the International Classification of Functioning, Disability and Health model from the World Health Organization [16] and how to compensate for disabilities; and (3) the perspective of dealing with the psychosocial consequences for the person with dementia and his family as described by the adaption-coping model of Dröes et al [17]. The content was composed by internationally recognized dementia experts. The platform aims to provide opportunities for collaboration, discussion, and sharing experiences between users across the EU. The main focus was to provide relevant content that was easy to find. Additionally, STAR aimed to promote accessibility to

specialized knowledge by experts in the field and to offer an online community of caregivers and other stakeholders.

To this end, the STAR training portal was designed and developed (Figure 1) to offer the following functionalities:

1. A collection of 8 modules on different topics in dementia care: 2 at a basic level and 6 at an intermediate and advanced level (Figures 2 and 3);
2. A Learning Path Advisor through an online tool integrated in STAR that assesses baseline knowledge and confidence to help people decide at which point to start the course; and
3. Facebook and LinkedIn communities to promote peer support and provide opportunities to contact other dementia care professionals.

The developed course is currently available for a nominal fee. It is fully available in English and Dutch with translation into Swedish, Italian, and Romanian underway at the time of writing this paper.

After development and testing of the training portal and e-learning course material during the first phase of the STAR project, the STAR training portal was evaluated in a randomized controlled trial (RCT) in the Netherlands and the United Kingdom from May 2013 to March 2014. The primary aim of this evaluation was to assess STAR's usefulness, user friendliness, and impact on knowledge. Because the themes of the course, in addition to factual knowledge on the dementia syndrome, focus greatly on dealing with dementia and understanding dementia (eg, themes such as "adaptation and coping," "positive and empathic communication," and "emotional impact and looking after yourself as a caregiver"), the impact on empathy, attitudes, and sense of competence were studied as well. The aim of this paper is to describe the results of these different types of user evaluations.

Figure 1. The main project page.

Figure 2. Overview of modules.

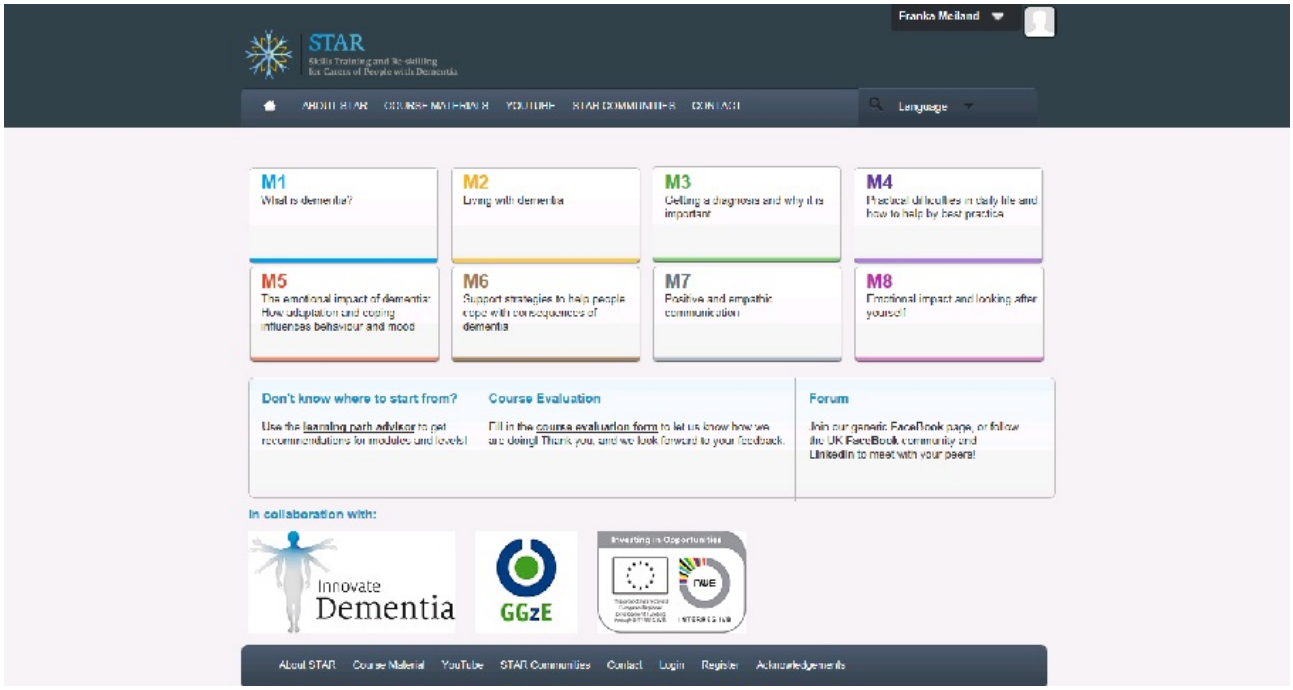
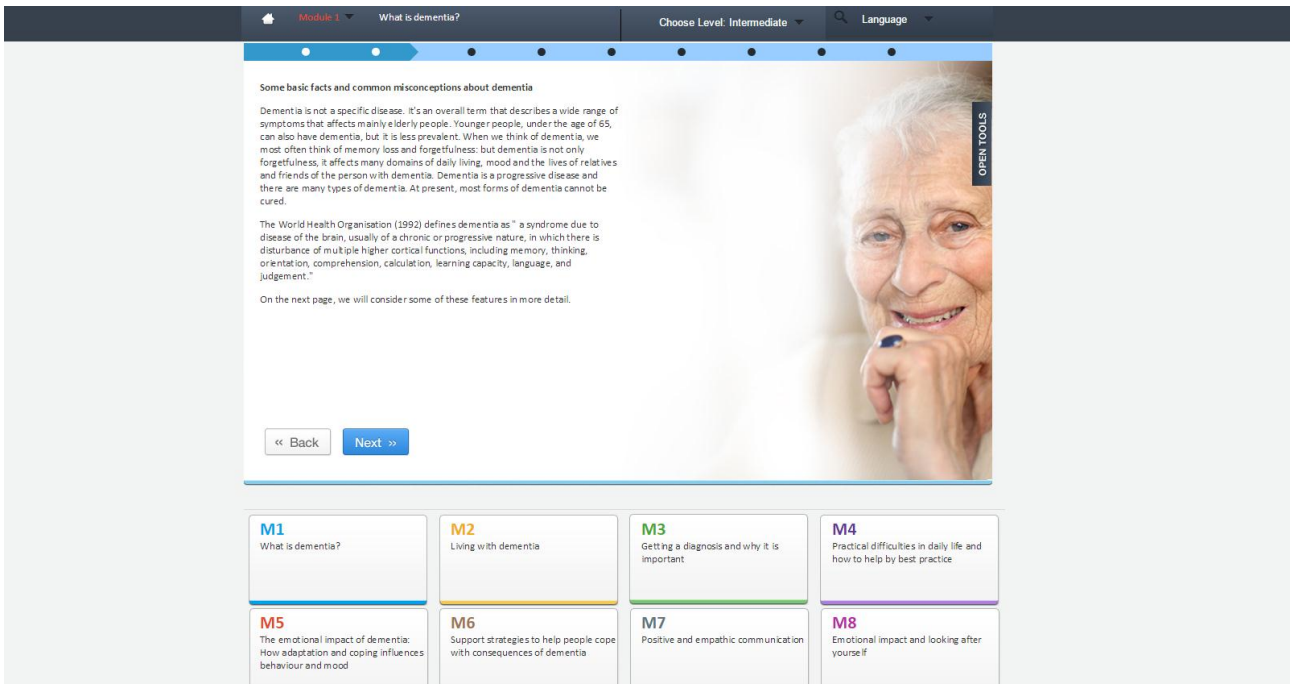


Figure 3. Example of a module (module 1: What is dementia?).



Methods

Design

An RCT design was used to assess the effectiveness of STAR among Dutch and English users. Participants were randomly assigned to either a group that could participate directly in the STAR training or to a group that had to wait for 4 months before they could register (free of charge) for the STAR training. Participants followed the course at their own pace; however, within a specified period of 4 months. Pretest data were gathered and follow-up data were collected after 2 to 4 months of finalizing the course in the experimental condition and after the same period in the waiting list group (control condition).

Setting and Participants

Participants were caring for someone with dementia as an informal caregiver, a volunteer in dementia care, or a professional caregiver, and were living in either the Netherlands or in the United Kingdom. Participants in the Netherlands were recruited through meeting centers for people with dementia and their caregivers, regional branches of the national Alzheimer's organizations, case managers, care organizations, and via announcements through several informative websites targeted at informal caregivers, volunteers, and those with an interest in

dementia. In the United Kingdom, participants were recruited through caregivers' cafes, church groups, university service users and caregiver groups, and local dementia care and welfare organizations.

Because STAR was developed both for informal caregivers (family caregivers and volunteers) as well as professional caregivers, participants who fulfilled the following criteria were recruited for the evaluation study: (1) were sufficiently computer literate to utilize the STAR website and (2) were currently an informal caregiver for someone with dementia living in the community, or a volunteer working with people with dementia with direct contact with community-dwelling people with dementia, or a professional caregiver for people with dementia with direct contact with community-dwelling people with dementia.

The STAR Training Portal

The STAR platform was designed to be accessible through any Internet-enabled device so users could access the course at any time and place of their convenience. The STAR training portal consists of an online course with 8 modules relating to different topics. These topics were selected to cover a wide range of topics relating to dementia and dementia care. The modules consist of text, videos, interactive exercises, knowledge tests, and also include references to other websites, literature, and videos.

The themes covered in the modules are as follows:

1. What is dementia?
2. Living with dementia
3. Getting a diagnosis and why it is important
4. Practical difficulties in daily life and how to help by best practice
5. The emotional impact of dementia: how adaptation and coping influences behavior and mood
6. Support strategies to help people cope with consequences of dementia
7. Positive and empathic communication
8. Emotional impact and looking after yourself

By answering questions from an interactive "learning advisor," participants are provided with advice relating to which module and level to start with so they may follow what may be considered a personalized learning path through the modules. For example, professionals with earlier experience in dementia care could be directed to the advanced levels of the course (including modules such as "Practical difficulties in daily life and how to help by best practice"), whereas informal caregivers who have never had to deal with dementia will be suggested to start their course with the basic modules on "What is dementia?" and "Living with dementia." Caregivers can then follow their own learning path, either gradually working their way up to the more advanced levels or they can choose not to progress beyond the basic modules/knowledge.

To ascertain what participants learn from the modules and to make the content more appealing, interactive exercises are included in the modules. These are used after each module at the basic and intermediate levels as quizzes to test level of knowledge. If an insufficient score is achieved in the quiz, participants are encouraged to reread the material and to try the quiz again following further learning.

Participants randomly allocated to the STAR training group could follow all STAR modules and were invited to take part in their national community (communities were created for all nationalities of users) on Facebook. They were explicitly asked to follow at least 4 modules, take part in the knowledge tests offered at the end of the modules, to complete the interactive exercises during the modules, and to watch a selection of videos that were offered in the modules.

Measurement Instruments

All questionnaires were offered online and were self-assessed in the participants' own language. Background characteristics were inventoried for all participants. These were age, sex, relation to the person with dementia (in case of informal caregivers), time involved in care for the person with dementia, and prior experience with courses on dementia.

For assessing usefulness and user friendliness, a questionnaire was composed specifically for this study, based on the Usefulness, Satisfaction, and Ease of use (USE) questionnaire [18] and online course evaluations used by the site programmers (AcrossLimits, Malta). This questionnaire contained 29 questions with 2 open questions, 20 other questions that could be answered on a 5-point scale ranging from "strongly agree" to "strongly disagree" (eg, "I instantly knew where to click"), and 7 questions on usefulness in which participants rated the usefulness of specific parts of STAR on a 4-point scale from "very useful" to "useless." Also, users were asked to indicate which modules they had followed and to grade each module on usefulness (1-10) to account for the fact that not all participants may have followed all modules of the course.

The primary outcome measures were knowledge on dementia and attitudes regarding dementia. Knowledge was measured with the Alzheimer's Disease Knowledge Scale (ADKS) [19] (internal consistency α=.71, test-retest reliability=0.81). The ADKS consisted of 30 questions on different aspects of Alzheimer's disease that could be answered with "true" or "false" (range 0-30), such as "People in their thirties can have Alzheimer's disease."

Attitudes toward dementia were assessed with 2 questions from the Alzheimer's disease survey [20] and approaches to dementia with the Approaches to Dementia Questionnaire (ADQ) [21] (α=.85 for person-centered care scale; α=.76 for hope scale; α=.83 total score). The latter questionnaire was also administered among informal caregivers with one question omitted ("It is important not to become too attached to people with dementia") because it was deemed inappropriate. The ADQ consisted of 19 questions on attitudes toward dementia and could be answered on a 5-point scale ranging from "completely agree" to "completely disagree" (range 19-95), such as "People with dementia are like children."

The secondary outcome measures were empathy, quality of life, burden, and sense of competence. The latter 3 were only administered among informal caregivers. Empathy was assessed with the Interpersonal Reactivity Index (IRI) [22]. This questionnaire consists of 28 items that were answered on a

5-point scale ranging from “does not describe me well” to “describes me very well” and with 4 subscales: (1) perspective taking (tendency to adopt the psychological point of view of others), (2) fantasy (tendency to imagine oneself into fictitious characters in books and movies), (3) empathic concern (“other-oriented” feelings of sympathy and concern for unfortunate others), and (4) personal distress (“self-oriented” feelings of anxiety and unease in tense interpersonal settings). The range was 0 to 28 for each subscale.

Quality of life was assessed with 2 distinct questions (eg, “how would you rate your quality of life on a scale from 1 to 10?”) and burden was assessed with 1 question. Finally, for sense of competence, the Short Sense of Competence Questionnaire (SSCQ; α=.77) [23] was used. The total score for the SSCQ was calculated by dichotomizing answers to 7 questions (eg, “I feel strained in my interactions with my...”) on a 5-point scale, counting only values of 4 or 5 (range 0-7).

Procedure

Participants, including informal caregivers (72/142, 50.7%), volunteers (24/142, 16.9%), and professional caregivers (46/142, 32.4%), in the Netherlands and United Kingdom were recruited through different partners (refer to Setting and Participants), both in person and through email. When people were interested in participating, a researcher provided them with additional written and oral information and a consent form. When a signed informed consent form was returned, the participants received a link to the online baseline questionnaire by email.

After having filled in the questionnaires, participants in the Netherlands and the United Kingdom were randomized to either the experimental or the control group. Participants were randomized based on the following variables. In each country, strata for each participant group—informal caregiver, volunteer, and professional—and within these strata (1) for informal caregivers, spouse of a person with dementia or not and knowledge regarding dementia being low (ADKS score <19), average (ADKS 20-26), or high (ADKS >27), and (2) for volunteers, shorter or longer than half a year of work experience and, for professionals, education level high or low.

Randomization software [24] was used to classify participants into either the experimental or control group. Participants in the experimental group received a link to the STAR registration webpage. People were free to choose the number of modules they followed with a baseline minimum of at least 4 to obtain a good impression of the course. People in the control group were informed that they were assigned to the group that could follow the course free of charge after post-test measurements 4 months later. At the end of the project, 2 to 4 months after the baseline measurement, all participants received a link to the questionnaires for post-test measurement. All personal data collected were anonymized. Participants were allocated a code number that was retained in a secured database under supervision of the project leaders at the evaluation sites.

Analyses

Descriptive analyses were performed to describe the baseline characteristics of the study population. Differences between the experimental group and the control group at baseline were analyzed with relevant difference tests (chi-square and *t* tests).

The usefulness and user friendliness of the STAR training were analyzed with descriptive statistics. Impact on the outcome measures was assessed with univariate covariance analyses (ANCOVAs) on the post-test data of the participants at 4 months, whereas pretest data were included as covariates. The background variables with baseline values that differed significantly between the experimental and control group and appeared to be related to one or more of the outcome measures (ie, potential confounding variables) were also included in the analyses as covariates.

Results

Description of Participants

In total, 142 persons participated in the STAR evaluation study. In the Netherlands, 85 people took part in the research. Of these, 50 persons were informal caregivers, 7 were volunteers in dementia care, and 28 were professional caregivers. In the United Kingdom, 57 people participated; 22 were informal caregivers, 17 were volunteers in dementia care, and 18 were professional caregivers. We grouped the informal caregivers and volunteers together as laypeople because of the relatively small number of volunteers. The background characteristics of the participants that completed both pretest and post-test measurements are detailed in Table 1.

Table 1. Background characteristics of participants at baseline for laypeople (informal caregivers and volunteers) and professional caregivers that finished pretest and post-tests.

Characteristic	Experimental group	Control group	*F* (*df1*,*df2*)	χ^2 (*df*)	*P*
Laypeople	n=27	n=32			
Age (years), mean (SD)	52.93 (11.43)	54.69 (14.36)	2.02 (1,57)		.16
Sex, n (%)				0.2 (58)	.65
Male	7 (26)	10 (31)			
Female	20 (74)	22 (69)			
Relationship, n (%)				4.0 (58)	.41
Partner	9 (33)	9 (28)			
Child	8 (30)	5 (16)			
Sister/brother	0 (0)	1 (3)			
Other	4 (15)	10 (30)			
NA	6 (22)	7 (22)			
Duration of care, n (%)				7.3 (58)	.17
<3 months	2 (7)	6 (19)			
3-12 months	2 (7)	1 (3)			
1-2 years	2 (7)	9 (28)			
2-5 years	15 (58)	12 (38)			
>5 years	6 (21)	4 (12)			
ADKS score,[a] mean (SD)	24.67 (3.43)	24.13 (3.32)	0.01 (1,57)		.92
Professionals	n=10	n=14			
Age (years), mean (SD)	46.90 (12.12)	48.07 (9.11)	1.04 (1,22)		.32
Sex, n (%)				3.1 (23)	.08
Male	2 (20)	0 (0)			
Female	8 (80)	14 (100)			
Duration of care, n (%)				3.5 (23)	.62
<3 months	1 (10)	2 (14)			
3-12months	1 (10)	4 (29)			
1-2 years	1 (10)	2 (14)			
2-5 years	3 (30)	1 (7)			
>5 years	4 (40)	5 (36)			
ADKS score,[a] mean (SD)	23.60 (3.40)	24.36 (3.52)	0.15 (1,22)		.70

[a]ADKS: Alzheimer's Disease Knowledge Scale.

During the pilot, 59 participants dropped out. The total response at post-test was 61%. Reasons for dropouts in the Netherlands (n=29) were no time (n=4) or unknown (n=25; no response to repeated emails of researchers to remind them of filling in the questionnaires). Reasons for dropouts in the United Kingdom (n=30) were no time (n=1), no computer at home (n=1), or unknown (n=28; no response to repeated requests by researchers to fill in the questionnaires). Due to a technical issue (the rule forcing participants to fill in all usefulness and user friendliness questions before continuing did not function), one Dutch participant did not fill in the questions on usefulness and user friendliness, although he filled in all impact questions. Analyses to test differences in background characteristics between completers and dropouts indicated that for both formal and informal caregivers and volunteers there were no significant differences in age, gender, relationship, and duration of care/work between these groups. Furthermore, at baseline there were no statistically significant differences in background characteristics and primary outcome measures, such as knowledge, empathy, and approaches between the experimental and control group. For an overview of participants, refer to the

flowchart in Multimedia Appendix 1. The CONSORT checklist for this study is shown in Multimedia Appendix 2.

Results of the Evaluation of Usefulness and User Friendliness of the STAR Training

At post-test, participants from the experimental group (following the STAR training) were asked to indicate if they followed a particular module and, if so, to rate its usefulness on a scale from 0 to 10. These ratings are presented in Table 2. Participants in the Netherlands and the United Kingdom were positive overall about the usefulness of the different modules. On average across the countries, the modules that were assessed as most useful were modules 4 (practical difficulties in daily life and how to help) and 6 (support strategies to help people cope with the consequences of dementia). The modules considered least useful were modules 1 (what is dementia?) and 3 (getting a diagnosis and why it is important).

Table 2. Rating of usefulness of the modules by the participants at post-test (mean score on scale 1-10).[a]

Module	Laypeople, mean (SD), n		Professionals, mean (SD), n	
	Netherlands n=17	UK n=9	Netherlands n=8	UK n=2
Baseline				
1. What is dementia?	8.07 (1.03), 15	8.22 (1.92), 9	8.40 (1.61), 7	6.00 (—), 1
2 Living with dementia	8.13 (1.13), 15	9.00 (0.76), 8	8.30 (1.80), 7	—
Intermediate				
3. Getting a diagnosis	7.76 (1.35), 17	8.88 (0.84), 8	7.75 (1.75), 8	—
4. Practical difficulties	7.76 (1.30), 17	9.14 (0.69), 7	7.55 (1.75), 7	—
5. Emotional impact of dementia	7.76 (1.56), 17	9.00 (0.82), 7	7.80 (1.67), 6	7.00 (—), 1
6. Support strategies	7.71 (1.45), 17	8.57 (0.98), 7	8.40 (1.51), 5	9.00 (—), 1
7. Empathic communication	7.53 (1.40), 15	9.00 (1.00), 5	8.00 (1.41), 4	10.00 (—), 1
8. Emotional impact for caregiver	7.86 (1,41), 14	9.00 (1.00), 5	8.00 (1.41), 4	10.00 (—), 1
Advanced				
3. Getting a diagnosis	7.71 (0.76), 7	6.40 (3,64), 5	6.83 (1.17), 6	—
4. Practical difficulties	7.63 (1.19), 8	8.00 (1.00), 3	6.00 (1.00), 3	—
5. Emotional impact of dementia	7.75 (1.04), 8	7.67 (0.58), 3	8.33 (1.53), 3	10.00 (—), 1
6. Support strategies	7.29 (0.76), 7	8.00 (1.00), 3	7.50 (0.71), 2	—
7. Empathic communication	7.50 (0.84), 6	7.50 (0.71), 2	7.50 (0.71), 2	—
8. Emotional impact for caregiver	7.88 (0.84), 8	7.50 (0.71), 2	7.67, 3	—
Mean overall rating	7.74 (0.87)	8.27 (0.41)	7.16 (0.66)	8.67 (—)

[a]A higher score means participants considered it to be more useful.

The results on the opinions about usefulness of the different elements, (eg, text or videos) of the STAR training are shown in Table 3. Opinions on user friendliness are shown in Table 4.

Table 3. Ratings on usefulness of specific elements of the STAR training (range 1-4).[a]

Element	Laypeople, median (interquartile range)		Professionals, median (interquartile range)	
	Netherlands n=17	UK n=9	Netherlands n=8	UK n=2
The text of the modules	3.0 (0.0)	4.0 (1.0)	3.0 (1.0)	3.5 (—)
The interactive exercises	3.0 (1.0)	4.0 (1.0)	3.0 (0.0)	3.5 (—)
The knowledge questions	3.0 (1.0)	4.0 (1.0)	3.0 (1.0)	3.5 (—)
The videos	4.0 (1.0)	2.0 (2.0)	3.0 (1.0)	3.0 (—)
The online community[b]	3.0 (1.0)	3.0 (—)	4.5 (2.0)	3.0 (—)
The comments from the expert community[c]	3.0 (1.0)	3.0 (0.0)	5.0 (2.0)	3.0 (—)

[a]Scoring: 1=useless; 2=a little useful; 3=useful; 4=very useful; NA=not applicable.

[b]N/A in NL: n=8; N/A in UK: n=7.

[c]N/A in NL: n=7; N/A in UK: n=2.

Table 4. Opinions on user friendliness (range 1-5).[a,b]

Question	Laypeople, median (interquartile range)		Professionals, median (interquartile range)	
	Netherlands n=17	UK n=9	Netherlands n=8	UK n=2
Logging in was easy	1.0 (0.0)	1.0 (0.0)	1.0 (0.0)	2.0 (—)
I immediately noticed where I have to click	1.0 (1.0)	1.0 (0.0)	1.5 (1.0)	2.0 (—)
The overall layout is simple to follow	1.0 (1.0)	1.0 (0.0)	1.0 (1.0)	1.5 (—)
The STAR training was easy to do	1.0 (1.0)	1.0 (0.0)	1.0 (1.0)	1.5 (—)
The material has been well thought out	1.0 (1.0)	1.0 (1.0)	1.0 (0.0)	1.0 (—)
The length of the modules and exercises was just right	1.0 (1.0)	1.0 (1.0)	1.0 (1.0)	1.0 (—)
The course was nice to do	1.0 (0.0)	1.0 (1.0)	1.0 (0.0)	1.5 (—)
I knew what I had to do in the STAR training (navigating, exercises, etc)	1.0 (0.0)	1.0 (0.0)	1.0 (1.0)	1.0 (—)

[a]Scoring: 1=completely agree; 2=agree a little; 3=agree/disagree; 4=disagree a little; 5=completely disagree.

[b]In the UK cases where there's only 2 participants, the interquartile range could not be computed.

Use of the Learning Path Advisor

Analysis of logging files of the STAR training indicated that most of the STAR participants in the experimental group used the Learning Path Advisor. In the Netherlands, 73% of informal caregivers and 91% of professional caregivers used the Learning Path Advisor to obtain a personalized suggestion where to start in the course. In the United Kingdom, 9% of informal caregivers and 17% of professionals used the Learning Path Advisor.

Impact on Outcome Measures

In Table 5 are the results of the ANCOVA analysis to assess the impact of STAR training compared to a waiting list condition on the primary and secondary outcome measures of participants in the Netherlands and the United Kingdom.

Table 5. Impact of STAR training on outcome measures for laypeople and professionals (Netherlands and United Kingdom together).

Outcomes[a]	Pretest, mean (SD)		Post-test, mean (SD)		*F* (*df1*,*df2*)	*P*	η^{2b}
	Experimental	Control	Experimental	Control			
Primary outcomes							
ADQ [c]							
Laypeople	(n=27)	(n=32)	(n=27)	(n=32)			
Total (score 18-90)	69.15 (6.74)	60.13 (10.4)	71.59 (6.48)	64.66 (4.90)	12.98 (1,57)	.001	0.19
Hope scale (score 8-40)	20.48 (4.26)	18.25 (3.89)	22.33 (5.33)	19.13 (3.68)	3.54 (1,57)	.07	0.06
Person scale (score 10-50)	48.67 (4.45)	41.87 (9.50)	49.26 (3.49)	45.53 (3.56)	7.48 (1,57)	.008	0.12
Professionals	(n=10)	(n=14)	(n=10)	(n=14)			
Total (score 19-95)	76.30 (5.42)	56.78 (19.6)	77.70 (5.42)	77.21 (8.12)	2.32 (1,22)	.14	0.10
Hope scale (score 8-40)	27.50 (3.89)	21.21 (5.74)	27.80 (3.68)	27.86 (4.59)	0.31 (1,22)	.59	0.01
Person scale (score 11-55)[d]	48.80 (2.78)	35.57 (14.4)	49.90 (3.03)	49.36 (4.41)	1.23 (1,22)	.28	0.06
ADKS [e] **(score 1-30)**							
Laypeople	(n=27)	(n=32)	(n=27)	(n=32)			
Total	24.67 (3.43)	24.13 (3.32)	24.44 (3.11)	24.28 (3.12)	0.02 (1,57)	.90	0.00
Professionals	(n=10)	(n=14)	(n=10)	(n=14)			
Total	23.60 (3.41)	24.36 (3.52)	24.20 (2.57)	24.64 (2.40)	0.00 (1,22)	.97	0.00
Attitudes (score 1-7)							
Laypeople	(n=24)	(n=30)	(n=24)	(n=30)			
Total	2.91 (1.78)	2.59 (1.82)	2.75 (1.85)	2.10 (1.67)	1.74 (1,52)	.19	0.03
Professionals	(n=2)	(n=4)	(n=2)	(n=4)			
Total	3.33 (1.32)	2.67 (1.61)	3.22 (1.39)	2.92 (1.68)	0.01 (1,4)	.93	0.00
Secondary outcomes							
Empathy (score 0-28)							
Laypeople	(n=27)	(n=32)	(n=27)	(n=32)			
Distress	14.33 (6.20)	14.25 (5.85)	9.74 (5.33)	13.59 (5.63)	9.89 (1,22)	.003	0.15
Empathy	12.56 (6.45)	12.81 (6.60)	20.40 (4.06)	13.03 (5.63)	47.63 (1,22)	<.001	0.46
Fantasy	13.70 (5.19)	13.75 (4.24)	14.30 (5.24)	12.84 (4.43)	1.41 (1,22)	.24	0.03
Perspective	13.11 (5.66)	13.06 (5.79)	18.81 (3.45)	13.75 (4.45)	25.90 (1,22)	<.001	0.32
Professionals	(n=10)	(n=14)	(n=10)	(n=14)			
Distress	13.30 (7.69)	13.86 (7.68)	7.50 (2.80)	14.57 (7.00)	17.95 (1,57)	<.001	0.46
Empathy	12.90 (6.56)	14.85 (8.59)	20.30 (2.50)	14.15 (8.12)	19.37 (1,57)	<.001	0.49
Fantasy	13.20 (3.49)	13.57 (4.18)	12.50 (4.35)	14.64 (5.37)	1.11 (1,57)	.23	0.05
Perspective	13.30 (5.74)	12.93 (6.06)	19.10 (3.21)	13.93 (7.43)	6.58 (1,57)	.02	0.24
Quality of life [f] **(grade 1-10)**	(n=21)	(n=25)	(n=21)	(n=25)			
Informal caregivers	7.24 (1.58)	6.23 (1.75)	7.05 (1.77)	6.48 (1.58)	0.00	.97	0.00
Burden [f] **(score 1-5)**							
Informal caregivers	2.67 (1.11)	3.08 (1.13)	2.43 (0.98)	2.80 (0.96)	0.63	.43	0.02

Outcomes[a]	Pretest, mean (SD)		Post-test, mean (SD)		*F* (*df1*,*df2*)	*P*	η^{2b}
	Experimental	Control	Experimental	Control			
Sense of competence (SSCQ)[f,g] (score 0-7)							
Informal caregivers	4.43 (1.25)	4.54 (1.56)	4.67 (1.06)	4.04 (1.49)	5.50	.02	0.11

[a]For all scores except distress, burden, and sense of competence, the higher score is the more favorable.

[b]Effect size (η^2) is considered small at 0.01, medium at 0.06, and large at 0.14.

[c]ADQ: Approaches to Dementia Questionnaire

[d]For professional caregivers, score is 11-55 because they had one additional question compared to informal caregivers.

[e]ADKS: Alzheimer's Disease Knowledge Scale

[f]These were questionnaires specifically relevant to informal caregivers; they were not applied in volunteers.

[g]SSCQ: Short Sense of Competence Questionnaire

For the primary outcome measures, we found a statistically significant effect on attitudes toward dementia (ADQ total score) with a large effect size (η^2) and on its subscale, person-centered care among laypeople (informal caregivers and volunteers), with a large effect size. For both cases, both the experimental group and the control group increased in score. Differences in baseline scores were accounted for by adding these as a covariate in the analyses. We did not find outliers at baseline that could explain this result. For the remaining primary outcomes (knowledge about and attitudes toward dementia), we did not find statistically significant differences in the experimental or control group among laypeople and professionals.

There were 2 separate questions relating to attitudes toward dementia. The first question related to which disease people were most afraid of developing (diabetes, stroke, heart disease, dementia, cancer, depression, or influenza). At pretest, 31% of laypeople were most afraid of developing dementia; for professionals, this was 13%. At post-test, 46% of laypeople and 8% of professionals were most afraid of developing dementia. This indicates that after completing the course, laypersons were more afraid of developing dementia, whereas this fear decreased in professionals. The second question on attitudes toward dementia was "if you had a family member who was exhibiting confusion and memory loss, would you want the person to see a doctor to determine if the cause of the symptoms was Alzheimer's disease or not?". This was answered positively by all who completed the research, both at pretest and at post-test.

For the secondary outcome measures, we found statistically significant differences between the experimental and control group in the expected direction on several subscales of empathy: the distress subscale, empathy subscale, and perspective subscale. Although the scores on these scales remained largely the same in the control group at pretest and post-test, we witnessed a significant improvement in the experimental group, indicating that they felt less distressed in tense situations, had more empathy and concern for the well-being of other people, and were better able to understand situations and the actions of other people. On the other hand, we found a medium-sized, significant negative effect on sense of competence of informal caregivers, which declined in the experimental group compared to the control group, implying that participants in the course felt less competent to fulfill their care task after following the course. Post hoc analysis revealed that this decline in sense of competence was related to a higher age of the caregivers (r=–.34, P=.02).

Discussion

The STAR training, an online e-learning course developed to skill/reskill informal caregivers, volunteers, and professional caregivers of persons with dementia, was evaluated in the Netherlands and the United Kingdom on its usefulness, user friendliness, and effectiveness. This evaluation was undertaken in an RCT, comparing participants who followed the STAR training to a waiting list control group. The evaluation results indicated that, in general, the 8 modules of the STAR training were positively/very positively valued with regard to usefulness and user friendliness. The course was considered easy to undertake and the material was considered well thought out. Participants indicated that the course made them feel more secure about their quality as a caregiver. Although all modules were assessed positively, some modules, such as modules 1 and 3, scored lower than others did. The content of these modules will need to be reviewed for future versions of STAR.

The results of the RCT in the Netherlands and the United Kingdom demonstrated a significant positive impact of the STAR training course on maintaining feelings of empathy among informal caregivers and volunteers. Also, an effect was found on a person-centered care approach; both the person-centered care approach and the total score on positive approaches toward dementia increased among laypeople in both the experimental and the control group. The sense of competence declined slightly in the informal caregivers who followed the course, which appeared to be related to the higher age of the caregivers. The decline in sense of competence is in contrast to the finding that a large number of the participants who followed the STAR training indicated that they felt more secure about their qualities as caregivers.

For professional caregivers, empathy improved among those who followed the course. This is an indication that after following the course they became better able to view situations from another's perspective (eg, a person with dementia) and that they showed more sympathy and concern, which may help them to provide better care for people with dementia. No effects

were found on knowledge about Alzheimer's disease. This was likely because the selected instrument, the ADKS, mostly has factual questions on symptoms and prevalence, although the STAR course predominantly focused on informing caregivers how to deal and cope with the consequences of dementia. Therefore, positive outcomes related to dealing with dementia (eg, empathy and attitudes), rather than an increase in knowledge, were in-line with our expectations. On other outcome measures, such as quality of life and burden, no effects were found.

These results are in-line with recent research, which has also found beneficial effects from Internet-based interventions [7], among others on attitudes and empathy [10]. One of the explanations for the effectiveness of STAR is the opportunity to choose "personalized learning paths," which make it possible for each individual caregiver to tailor the content of the training to their own needs and skill level by recommending which modules are most relevant to them. According to earlier research by Lustria et al [8], tailored computer-based health interventions lead to improved health in caregivers. Previous research also indicates that increasing empathy in caregivers is highly relevant because it increases the well-being of the person with dementia [25].

Review studies on the effectiveness of psychosocial and technology-driven interventions to support family caregivers show that interventions using a psychoeducational or psychotherapeutic approach appear to be among the most powerful psychosocial interventions to improve quality of life of persons with dementia and their caregivers, and delay patient institutionalization [26-28]. Nevertheless, many studies suffered from serious methodological problems, such as unclear randomization methods, inadequate power calculation, selectively reported outcomes, and no use of an intention-to-treat analysis [29-31]. In addition, interventions were difficult to compare because type and intensity varied [32]. STAR adds to this research by offering an RCT with clear randomization methods and adequate numbers for sufficient statistical power. Additionally, all outcomes of STAR are clearly reported.

A strength of this study is that the STAR training portal was tested in an RCT design in 2 countries. A limitation of the study was the high number of dropouts in the RCT (34% in the Netherlands, 53% in the United Kingdom), especially in the experimental group; 43% (Netherlands) and 64% (United Kingdom) dropped out of the study. It seems likely that participants in the control group were more motivated to participate in the post-test because participating in the questionnaires would offer them free access to the course afterward. Furthermore, given that the STAR training portal was still in late development at the time of testing, some errors occurred when people followed the course. Another limitation was the fact that the online communities were not used often by participants, making them less informative and supportive than originally anticipated. The communities only contained a small number of participants because access was limited to only those in the experimental group of the research. However, they are expected to become more active and supportive in the future. This expectation is based on the fact that the STAR website has recently been updated to show links to the community websites more clearly. Additionally, when STAR gets more users, more users will potentially visit these communities, making them more lively and, therefore, more interesting to use for other visitors. We found little use for the Learning Path Advisor in the United Kingdom compared to use of it in the Netherlands. One explanation for this could be that the British group consisted of more volunteers with less experience in dementia care, who may have been more interested to follow the entire course, whereas the Dutch group consisted largely of informal caregivers and the majority (79%) had been caring for a person with dementia for 2 years or longer. It is likely that these experienced caregivers tended to use the Learning Path Advisor more frequently to find out which modules would provide them with new information, taking into account the knowledge they already had.

In conclusion, based on the promising results of our study, especially the positive effects of the STAR training portal on empathy for both laypeople (informal caregivers and volunteers) and professionals, it is recommended to repeat the RCT on a larger scale and in more countries. STAR is currently available in Dutch and English; the basic and intermediate modules are available in Italian and Romanian, and some are available in Swedish as well.

The positive effects of STAR on attitudes and empathy of caregivers may contribute to appropriate and high-quality dementia care in the community now and in the future. Therefore, an Internet-based intervention such as STAR can be a very useful alternative for face-to-face education and support for caregivers/informal caregivers and low-/unschooled professionals [2,3] and at lower costs [9], thus providing a means to cope with the challenge of taking care of the growing number of people with dementia in our society.

Acknowledgments

We gratefully acknowledge all participants in the RCTs in both the Netherlands and United Kingdom who participated in the course and supplied us with their opinions. We also gratefully acknowledge the writers of the course material. The STAR project was funded by the Leonardo da Vinci Life Long Learning Programme of the European Union (no. 510364-2010) and the BAVO Foundation in the Netherlands.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Participant flowchart for CONSORT.

[PDF File (Adobe PDF File), 9KB - jmir_v17i10e241_app1.pdf]

Multimedia Appendix 2

CONSORT checklist PDF.

[PDF File (Adobe PDF File), 9MB - jmir_v17i10e241_app2.pdf]

References

1. Prince M, Bryce R, Ferri C. World Alzheimer Day Report, Alzheimer's Disease International. London, UK: Alzheimer's Disease International (ADI); Sep 01, 2011.
2. Glueckauf R, Ketterson T, Loomis J, Dages P. Online support and education for dementia caregivers: overview, utilization, and initial program evaluation. Telemed J E Health 2004;10(2):223-232. [Medline: 15319052]
3. Martín-Carrasco M, Martín MF, Valero CP, Millán PR, García CI, Montalbán SR, et al. Effectiveness of a psychoeducational intervention program in the reduction of caregiver burden in Alzheimer's disease patients' caregivers. Int J Geriatr Psychiatry 2009 May;24(5):489-499. [doi: 10.1002/gps.2142] [Medline: 18949763]
4. Mitchell R. Planning for instructional technology in the classroom. New Directions for Community Colleges 2011 Jun 22;2011(154):45-52. [doi: 10.1002/cc.445]
5. Serafini JD, Damianakis T, Marziali E. Clinical practice standards and ethical issues applied to a virtual group intervention for spousal caregivers of people with Alzheimer's. Soc Work Health Care 2007;44(3):225-243. [doi: 10.1300/J010v44n03_07] [Medline: 17548277]
6. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of E-learning in medical education. Acad Med 2006 Mar;81(3):207-212. [Medline: 16501260]
7. Murray E, Burns J, See T, Lai R, Nazareth I. Interactive health communication applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
8. 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]
9. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
10. Irvine AB, Billow MB, Gates DM, Fitzwater EL, Seeley JR, Bourgeois M. Internet training to respond to aggressive resident behaviors. Gerontologist 2012 Feb;52(1):13-23 [FREE Full text] [doi: 10.1093/geront/gnr069] [Medline: 22038338]
11. Macdonald C, Stodel E, Chambers LW. An online interprofessional learning resource for physicians, pharmacists, nurse practitioners, and nurses in long-term care: benefits, barriers, and lessons learned. Inform Health Soc Care 2008 Mar;33(1):21-38. [doi: 10.1080/14639230801886824] [Medline: 18604760]
12. Van de Steeg L, Langelaan M, Ijkema R, Wagner C. The effect of a complementary e-learning course on implementation of a quality improvement project regarding care for elderly patients: a stepped wedge trial. Implement Sci 2012;7:13 [FREE Full text] [doi: 10.1186/1748-5908-7-13] [Medline: 22385554]
13. Van der Roest H. STAR: Skills Training and Re-skilling for Carers of People with Dementia. 2011. Deliverable D2.1: reporting on existing courses URL: http://www.startraining.eu/index.php/en/ [accessed 2015-10-13] [WebCite Cache ID 6cFZvivC7]
14. Hattink B. STAR: Skills Training and Re-skilling for Carers of People with Dementia. 2014. STAR training portal URL: http://courses.startraining.eu/ [accessed 2014-11-11] [WebCite Cache ID 6U0DrH9wm]
15. Diagnostic and Statistical Manual of Mental Disorders (4th edition). Arlington, VA: American Psychiatric Pub; 2000.
16. Jette AM. Toward a common language for function, disability, and health. Phys Ther 2006 May;86(5):726-734 [FREE Full text] [Medline: 16649895]
17. Dröes RM, Roest HG, Mierlo L, Meiland FJM. Memory problems in dementia: adaptation and coping strategies and psychosocial treatments. Expert Rev Neurother 2011 Dec;11(12):1769-81; quiz 1782. [doi: 10.1586/ern.11.167] [Medline: 22091600]
18. Lund A. Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter 2001 Jan;8(2):1.
19. Carpenter BD, Balsis S, Otilingam PG, Hanson PK, Gatz M. The Alzheimer's Disease Knowledge Scale: development and psychometric properties. Gerontologist 2009 Apr;49(2):236-247 [FREE Full text] [doi: 10.1093/geront/gnp023] [Medline: 19363018]
20. Blendon RJ, Benson JM, Wikler EM, Weldon KJ, Georges J, Baumgart M, et al. The impact of experience with a family member with Alzheimer's Disease on views about the disease across five countries. Int J Alzheimers Dis 2012;2012:903645 [FREE Full text] [doi: 10.1155/2012/903645] [Medline: 22997601]
21. Lintern T, Woods B, Phair L. Training is not enough to change care practice. J Dem Care 2000 Jan 01;8(2):15-17.

22. Davis M. Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of Personality and Social Psychology 1983;44(1):113-126. [doi: 10.1037/0022-3514.44.1.113]
23. Vernooij-Dassen MJ, Felling AJ, Brummelkamp E, Dauzenberg MG, van den Bos GA, Grol R. Assessment of caregiver's competence in dealing with the burden of caregiving for a dementia patient: a Short Sense of Competence Questionnaire (SSCQ) suitable for clinical practice. J Am Geriatr Soc 1999 Feb;47(2):256-257. [Medline: 9988301]
24. Saghaei M. Random allocation software for parallel group randomized trials. BMC Med Res Methodol 2004 Nov 9;4:26 [FREE Full text] [doi: 10.1186/1471-2288-4-26] [Medline: 15535880]
25. Norton M, Fauth E, Piercy K, Corcoran C, Hess K, Morrison A, et al. Higher caregiver agreeableness predicts slower cognitive decline in persons with Alzheimer's disease: The dementia progression study. Alzheimer's & Dementia 2010 Jul;6(4):S463. [doi: 10.1016/j.jalz.2010.05.1541]
26. Smits CH, Lange LJ, Dröes RM, Meiland FJ, Vernooij-Dassen M, Pot AM. Effects of combined intervention programmes for people with dementia living at home and their caregivers: a systematic review. Int J Geriatr Psychiatry 2007 Dec;22(12):1181-1193. [doi: 10.1002/gps.1805] [Medline: 17457793]
27. Moniz-Cook E, Vernooij-Dassen M, Woods B, Orrell M. Psychosocial interventions in dementia care research: the INTERDEM manifesto. Aging Ment Health 2011 Apr;15(3):283-290. [doi: 10.1080/13607863.2010.543665] [Medline: 21491215]
28. Dröes RM, Breebaart E, Meiland FJ, Tilburg W, Mellenbergh GJ. Effect of Meeting Centres Support Program on feelings of competence of family carers and delay of institutionalization of people with dementia. Aging Ment Health 2004 May;8(3):201-211. [doi: 10.1080/13607860410001669732] [Medline: 15203401]
29. Schoenmakers B, Buntinx F, DeLepeleire J. Supporting the dementia family caregiver: the effect of home care intervention on general well-being. Aging Ment Health 2010 Jan;14(1):44-56. [doi: 10.1080/13607860902845533] [Medline: 20155520]
30. Thompson C, Spilsbury K, Hall J, Birks Y, Barnes C, Adamson J. Systematic review of information and support interventions for caregivers of people with dementia. BMC Geriatr 2007;7:18 [FREE Full text] [doi: 10.1186/1471-2318-7-18] [Medline: 17662119]
31. Cooke D, McNally L, Mulligan K, Harrison MJ, Newman SP. Psychosocial interventions for caregivers of people with dementia: a systematic review. Aging Ment Health 2001 May;5(2):120-135. [doi: 10.1080/713650019] [Medline: 11511059]
32. Godwin K, Mills W, Anderson J, Kunik ME. Technology-driven interventions for caregivers of persons with dementia: a systematic review. Am J Alzheimers Dis Other Demen 2013 May;28(3):216-222. [doi: 10.1177/1533317513481091] [Medline: 23528881]

Abbreviations

ADKS: Alzheimer’s Disease Knowledge Scale
ADQ: Approaches to Dementia Questionnaire
DSM-IV: Diagnostics and Statistics Manual, IV (4th) Edition
EU: European Union
IRI: Interpersonal Reactivity Index
RCT: randomized controlled trial
SSCQ: Short Sense of Competence Questionnaire
STAR: Skills Training And Reskilling
USE: Usefulness, Satisfaction, and Ease of use

Edited by G Eysenbach; submitted 12.11.14; peer-reviewed by M Kazmer, B Fauth, M Chiu; comments to author 12.02.15; revised version received 07.04.15; accepted 21.06.15; published 30.10.15.

Please cite as:
Hattink B, Meiland F, van der Roest H, Kevern P, Abiuso F, Bengtsson J, Giuliano A, Duca A, Sanders J, Basnett F, Nugent C, Kingston P, Dröes RM
Web-Based STAR E-Learning Course Increases Empathy and Understanding in Dementia Caregivers: Results from a Randomized Controlled Trial in the Netherlands and the United Kingdom
J Med Internet Res 2015;17(10):e241
URL: http://www.jmir.org/2015/10/e241/
doi:10.2196/jmir.4025
PMID:26519106

Original Paper

Evaluating the Validity of an Automated Device for Asthma Monitoring for Adolescents: Correlational Design

Hyekyun Rhee[1], PhD; Michael J Belyea[2], PhD; Mark Sterling[3], PhD; Mark F Bocko[4], PhD

[1]School of Nursing, University of Rochester, Rochester, NY, United States

[2]College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[3]Department of Computer Science, School of Science and Technology, Nazarbayev University, Astana, Kazakhstan

[4]Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY, United States

Corresponding Author:
Hyekyun Rhee, PhD
School of Nursing
University of Rochester
601 Elmwood Ave, Box SON
Rochester, NY, 14642
United States
Phone: 1 585 276 3775
Fax: 1 585 273 1270
Email: hyekyun_rhee@urmc.rochester.edu

Abstract

Background: Symptom monitoring is a cornerstone of asthma self-management. Conventional methods of symptom monitoring have fallen short in producing objective data and eliciting patients' consistent adherence, particularly in teen patients. We have recently developed an Automated Device for Asthma Monitoring (ADAM) using a consumer mobile device as a platform to facilitate continuous and objective symptom monitoring in adolescents in vivo.

Objective: The objectives of the study were to evaluate the validity of the device using spirometer data, fractional exhaled nitric oxide (FeNO), existing measures of asthma symptoms/control and health care utilization data, and to examine the sensitivity and specificity of the device in discriminating asthma cases from nonasthma cases.

Methods: A total of 84 teens (42 teens with a current asthma diagnosis; 42 without asthma) aged between 13 and 17 years participated in the study. All participants used ADAM for 7 consecutive days during which participants with asthma completed an asthma diary two times a day. ADAM recorded the frequency of coughing for 24 hours throughout the 7-day trial. Pearson correlation and multiple regression were used to examine the relationships between ADAM data and asthma control, quality of life, and health care utilization at the time of the 7-day trial and 3 months later. A receiver operating characteristic (ROC) curve analysis was conducted to examine sensitivity and specificity based on the area under the curve (AUC) as an indicator of the device's capacity to discriminate between asthma versus nonasthma cases.

Results: ADAM data (cough counts) were negatively associated with forced expiratory volume in first second of expiration (FEV_1) (r=–.26, P=.05), forced vital capacity (FVC) (r=–.31, P=.02), and overall asthma control (r=–.41, P=.009) and positively associated with daily activity limitation (r=.46, P=.01), nighttime (r=.40, P=.02) and daytime symptoms (r=.38, P=.02), and health care utilization (r=.61, P<.001). Device data were also a significant predictor of asthma control (β=–.48, P=.003), quality of life (β=–.55, P=.001), and health care utilization (β=.74, P=.004) after 3 months. The ROC curve analysis for the presence of asthma diagnosis had an AUC of 0.71 (95% CI 0.58-0.84), which was significantly different from chance (χ^2_1=9.7, P=.002), indicating the device's discriminating capacity. The optimal cutoff value of the device was 0.56 with a sensitivity of 51.3% and a specificity of 72.7%.

Conclusions: This study demonstrates validity of ADAM as a symptom-monitoring device in teens with asthma. ADAM data reflect the current status of asthma control and predict asthma morbidity and quality of life for the near future. A monitoring device such as ADAM can increase patients' awareness of the patterns of cough for early detection of worsening asthma and has the potential for preventing serious and costly future consequences of asthma.

(J Med Internet Res 2015;17(10):e234) doi:10.2196/jmir.4975

KEYWORDS

asthma; adolescent; ambulatory monitoring; device; cough; validity

Introduction

Achieving acceptable asthma control in adolescents remains elusive despite the availability of efficacious treatment options. In 2010, nearly 11% of adolescents (2.7 million) aged 12-17 years in the United States reported a current diagnosis of asthma [1] and adolescents suffer greater asthma-related morbidity than other age groups [2]. Adverse asthma outcomes in this age group are largely attributable to poor self-management [3-5]. Establishing daily routines of symptom monitoring is recognized as the initial step to successful asthma self-management [6,7] leading to better asthma outcomes [8].

National Heart, Lung, and Blood Institute (NHLBI) Expert Panel Review 3 (EPR 3) [9] recommends that all patients with asthma learn to recognize symptom patterns related to inadequate asthma control. Currently, there are 2 basic types of home-based symptom-monitoring methods that patients can use to monitor symptom patterns: symptom-based and peak flow monitoring. However, many studies have consistently raised a concern about young people's poor perception in recognizing asthma symptoms [10-17]. Adolescents, in particular, tend to be overly optimistic in rating their asthma control despite the presence of symptoms and activity limitations [18-20]. As a consequence, sole reliance on patients' perception in symptom monitoring can be misleading.

Peak flow monitoring has been recommended to enhance the objectivity of symptom monitoring, particularly for those who suffer a high level of asthma severity because of their impaired symptom perception [21-23]. However, confirming the reliability of peak expiratory flow rate values has been an ongoing challenge [24-26]. Moreover, its clinical usefulness in children and adolescents is hampered by users' poor adherence and inadequate techniques, and the effort-dependent nature of the method. Concerns have been raised about questionable long-term sustainability of peak flow monitoring and inaccurate and/or fabricated readings [27-32].

Given the limitations of the existing symptom-monitoring methods, alternative strategies have been called for to mitigate the previously mentioned issues and enhance objectivity and sustainability for continuous symptom monitoring in children and adolescents [8,29,32,33]. Recently, we developed an Automated Device for Asthma Monitoring (ADAM) to increase the objectivity of symptom monitoring and to facilitate adolescents' adherence to continuous symptom monitoring in vivo. The device employed audio analysis technology to recognize symptoms, particularly coughs. The device uses a mobile system, iPod, as a platform. The methodological and technical details involving the development of the device and user acceptability are reported elsewhere [34,35]. This paper reports findings on the validity of the device as a monitoring tool. Specifically, this study examined (1) concurrent validity by correlating the data of the device with other measurements informing asthma control, including pulmonary function, fractional exhaled nitric oxide (FeNO), symptom-based monitoring (eg, daily asthma diary), asthma control questions, quality of life, and health care utilization; (2) predictive validity of the device by examining the extent to which the results of the device predict asthma control, quality of life, and health care utilization after 3 months; and (3) sensitivity and specificity of the device in discriminating between an asthma group and a nonasthma group.

Methods

Study Sample and Setting

Participant eligibility criteria for the asthma group were (1) age 13-17 years, (2) physician-diagnosed asthma for at least 1 year, and (3) ability to understand spoken and written English. The nonasthma group were age-matched adolescents with no current/past history of asthma and free of ongoing respiratory conditions. For both groups, we excluded those with other diagnoses producing respiratory symptoms (eg, upper respiratory infection, cardiac disease, cystic fibrosis) or significant cognitive impairment that could interfere with following the study protocol. Potential participants were recruited from the pediatric emergency department (ED) and outpatient clinics (primary practice and pediatric pulmonary practice) in a major university medical center located in the Northeastern United States. Of a total of 84 participants, most (73%, 61/84) were recruited from the ED and the remaining were from the study flyers (23%, 19/84) and clinician referrals from outpatient clinics (4%, 4/84). Unverifiable asthma diagnosis by medical records was the most common cause of ineligibility for the asthma group (n=52) followed by comorbidity with other respiratory diagnosis (n=14). For the nonasthma group, having an asthma diagnosis in the past (n=22) and presenting respiratory symptoms (n=7) were common reasons for ineligibility.

Study Measures: Both Groups

Automated Device for Asthma Monitoring

The ADAM device uses an iPod as a platform and was designed to continuously process audio data in real time to detect coughs. The device detected the number of cough events in 6-second intervals. It also provided a display of cough count data in a chart form on the device for users. Detailed descriptions of the device have been reported elsewhere [34,35]. ADAM was used by all participants for at least 7 consecutive days.

Fractional Exhaled Nitric Oxide

FeNO is a noninvasive method of assessing asthmatic inflammation [36]. Increasing FeNO levels have been found to be predictive of deteriorating asthma [37] and correlates more closely with symptoms than does forced expiratory volume in the first second (FEV_1) [38,39]. FeNO was measured before and after the 7-day trial in accordance with the American Thoracic Society (ATS) recommendations [40] using NIOX MINO (Aerocrine AB, Stockholm, Sweden).

Pulmonary Function Test

To assess the degree of airway obstruction, the volume of air expired during the first second of a forced vital capacity maneuver (FEV_1) and forced vital capacity (FVC) was measured using a KoKo spirometer (Pulmonary Data Service; Louisville, CO, USA) connected to a personal computer. Trained research staff performed spirometry for each participant two times, before and after the 7-day trial in accordance with the ATS standards [41].

Participant Demographic Form

Basic demographic information was collected, including gender, age, race, annual family income, years with asthma diagnosis (for the asthma group), types of health conditions that led to a clinic or ED visit (for the comparison group), and current medications (if applicable).

Study Measures: Asthma Group Only

Asthma Control Questions

The Asthma Control Questionnaire (ACQ) was developed based on the 2007 NHLBI National Guidelines' asthma control classification criteria involving 4 areas of asthma impairment, including the frequency of daily activity limitations, asthma symptoms, nighttime symptoms, and use of short-acting beta agonists (SABA) in the past 4 weeks. The 4 questions were measured on a 5-point scale and higher total scores indicated better asthma control. The ACQ was administered at pretrial and at 6-month follow-up. Cronbach alpha of the scale was .79 in this study.

Pediatric Asthma Quality of Life Questionnaire

The 23-item Pediatric Asthma Quality of Life Questionnaire (PAQLQ) measures 3 subdomains pertaining to asthma-related quality of life in children with asthma aged 7-17 years, including symptoms (10 items), emotional function (8 items), and activity limitation (5 items) [42]. Each item is measured on a 7-point scale (1=maximum impairment, 7=no impairment). Higher total scores indicate better levels of functioning. In this study, high internal consistency (Cronbach alpha) was found in all 3 subscales: .94, .95, and .88 for symptoms, emotional function, and activity limitation subdomains, respectively.

Health Care Utilization Form

The Health Care Utilization Form captured any asthma-related events including ED visits, hospitalization, office visits, and missed school days. This form assessed the frequency of events that occurred over the past 3 months (pretrial, 3-month follow-up) and the past 7 days for post-trial.

Visual Analog Scale

The Visual Analog Scale (VAS) is a line 100 mm long with 3 anchors dividing 3 zones (red, yellow, and green). For each symptom, there is a green zone (80-100 mm) labeled "no symptoms," a yellow zone (79-50 mm) labeled "mild symptoms," and a red zone (<50 mm) labeled "very bad symptoms" [16,29]. Teens marked any point on the line according to their perception of asthma symptoms two times a day in the morning and evening during the 7-day trial. The distance between the 0-mm mark and the placement of the "X" was measured to provide a numeric interpretation of their symptom perception.

Asthma Control Diary

The Asthma Control Diary consisted of 6 items, each with scores ranging from 0 (no symptoms) to 6 (continual symptoms) [43]. The device automatically sent diary reminders two times a day during the 7-day trial and allowed teens to conveniently complete the diary electronically using the touchscreen. The device automatically triggered diary reminders and made diary questions available only within the designated time window for am (6 am-noon) or pm (6 pm-midnight) to minimize recollection errors or the risk of data fabrication later. Morning questions pertained to nocturnal waking and morning symptoms, and evening questions assessed the degree of activity limitation, daytime symptoms (shortness of breath and wheeze), and SABA use in previous 24 hours. The mean score for each diary question was computed with higher scores indicating a greater degree of symptoms.

Study Procedure

At enrollment, spirometry and FeNO tests were conducted for all participants followed by the measurement of asthma control and health care utilization for the asthma group. All participants used the device continuously for the next 7 days during which the asthma group completed daily the electronic asthma diary and the VAS in the morning and at bedtime. On completion of the 7-day trial, spirometry and FeNO tests were repeated for all participants. In addition, the asthma group completed a quality of life questionnaire and reported any health care utilization that occurred in the past week. Follow-up data on asthma control, health care utilization, and quality of life were collected at 3 months after the trial only from the asthma group. Only 2 of 42 participants in the asthma group were lost to the 3-month follow-up. Of the 40 follow-up cases, 30 were completed by mail and 10 were in-person with research staff. This study protocol was approved by the Institutional Review Board, the Research Subjects Review Board, located in the University of Rochester Medical Center, Rochester, NY, and informed consent and assent were obtained from parents and teens, respectively. The participants received a monetary incentive (US $130 for the asthma group; US $100 for the nonasthma group) for their participation.

Data Analysis

All analyses were performed using SAS v9.3 (SAS Institute, Inc, Cary, NC, USA). Descriptive statistics were used to examine demographic and clinical characteristics. To assess concurrent validity, Pearson correlations were calculated between the device data from the asthma group with spirometer data (FEV_1 and FVC), FeNO, and other measures of asthma conditions, including asthma control questions, daily symptom diaries, VAS, quality of life, and health care utilization. To assess predictive validity of the device, regression was used to examine the relationships between the device data and asthma control, health care utilization, and quality of life collected at 3 months after the 7-day trial. Age, gender, and race were adjusted for in the regression analyses. Sensitivity and specificity were evaluated by assessing the device's capability to classify

2 distinctive groups. A receiver operating characteristic (ROC) curve analysis was conducted to calculate the area under the curve (AUC) as an indicator of the device's capacity to discriminate between asthma and nonasthma participants. An AUC of 1 indicates perfect classification and an AUC of 0.5 indicates that the ability to correctly classify is no better than chance. Based on the ROC, an optimum cutoff value was chosen to jointly maximize the sensitivity and specificity of the device. Given the few studies looking at devices for symptom monitoring in adolescents, any noticeable statistical trend is valuable; therefore, the significance level used in the analysis was less than .10. These findings are useful for an initial understanding of the validity of ADAM and for providing direction for further studies. Although it is understood that these findings will require further investigation, the risk of rejecting important research hypotheses was judged more important than the risk of type I error.

Results

Sample Characteristics

Details of participant flow from screening to the 3-month follow-up and sociodemographic characteristics of each group in the sample are reported elsewhere [34]. A total of 84 adolescents aged between 13 and 17 years (mean 15 years, SD 1.4) participated in the study. Of those, 61% (51/84) were females and 44% (37/84) were nonwhite adolescents. No significant differences were found between the asthma and nonasthma groups in age and gender. The asthma group included significantly more nonwhite adolescents, predominantly African American. At enrollment, the asthma group had slightly lower FEV_1 (mean 88.3%, SD 16.3 vs mean 90.9%, SD 13.8) and elevated FeNO (mean 28.6 ppb, SD 38.6 vs mean 25.6 ppb, SD 24.8) than the nonasthma group, yet these differences were not statistically significant. Within the asthma group, 19 of 42 (45%) reported active asthma symptoms at enrollment. Mean years since asthma diagnosis was 10.4 (SD 4.9) years. SABA use was reported by 95% (40/42) of the asthma group and most (60.5%, 25/42) were on at least one controller medication. The most common medication was inhaled corticosteroids (ICS), which was reported by 38% (16/42) of the asthma group, followed by ICS and long-term beta-agonist combination (29%, 12/42), and leukotriene modifier (21%, 9/42). No significant difference in control medication use was found between the symptomatic and nonsymptomatic groups. Only 2 participants (5%, 2/42) reported oral steroids as a current medication.

Coughs in the Asthma Group Monitored by the Device

Descriptive analysis was conducted on data from the asthma group excluding 3 asthma teens for which no data were recorded in the device due to unknown mechanical issues. All 39 teens with asthma used the device for a mean 8.26 (SD 1.47) consecutive days (median 8, range 5-14 days) and each teen used the device for a mean 19.4 hours/day (SD 1.71; median 19.7, range 15-22 hours). When the number of coughs was compared for morning (6 am-11:59 am), afternoon/evening (noon-10 pm), and bedtime (10 pm-6 am), a greater number of coughs were registered during afternoon/evening compared to morning and bedtime (P=.01 and P=.004, respectively). The number of coughs was not significantly different by gender or age.

Concurrent Validity

Cough Counts and Fractional Exhaled Nitric Oxide and Pulmonary Function

No significant correlations were found between the device's cough data and FeNO. The number of coughs was negatively associated with FEV_1 (r=-.26, P=.05) and FVC (r=-.31, P=.02) at enrollment indicating that the higher number of coughs was associated with poor pulmonary function.

Cough Counts and Daily Symptom Diary Data and Visual Analog Scale

Cough counts showed a positive association with limited activities (r=.46, P=.01) and shortness of breath (r=.29, P=.07). Cough counts showed significant associations with symptoms as measured by VAS at nighttime (r=.40, P=.02) and daytime (r=.38, P=.02).

Cough Counts and Asthma Control and Quality of Life

Table 1 shows correlations between cough counts and asthma control and quality of life during the trial. Cough counts were negatively associated with asthma control and overall quality of life, and the activity and symptom subscales of quality of life.

Table 1. Associations between the number of coughs and asthma control and quality of life.

Associated variables	r	P
Asthma control	-.41	.01
Quality of life total	-.28	.08
Activity subscale	-.27	.09
Symptoms subscale	-.29	.07
Emotional function subscale	-.26	.11

Cough Counts and Health Care Utilization

There was no association between cough counts and health care use in the 3 months before the trial. However, cough counts showed positive association with health care use during the 7-day trial (r=.61, P<.001) indicating that higher cough counts were associated with the higher use of health care services, particularly with the number of days of hospitalization (r=.72, P<.001) and office visits (r=.72, P<.001). Cough counts were

also positively associated with the number of missed school days (r=.70, P<.001).

Predictive Validity

Greater number of coughs was correlated with poor asthma control and quality of life and higher use of health care services 3 months later (Table 2).

Table 2. Correlations between the number of coughs and asthma control, quality of life, and health care utilization at 3 months after the 7-day trial.

Dependent variables	r	P
Asthma control	-.49	.002
Quality of life total	-.47	.004
Activity subscale	-.45	.006
Symptoms subscale	-.45	.006
Emotional function subscale	-.44	.007
Health care utilization	.55	.02

Table 3 presents the extent to which cough counts predicted asthma control, quality of life, and health care utilization after adjusting for age, gender, race, and family income. Coughs significantly predicted asthma control, and the regression model explained 42% of the variance in asthma control. Cough counts also significantly predicted the quality of life total score (β=-.55, P=.001) and each of the subscales including activity, symptoms, and emotional function 3 months later. Explained variance in quality of life was 38% for the quality of life total score, and 28%, 35%, and 41% for activity, symptoms, and emotional subscales, respectively. Health care utilization during the 3 months following the 7-day trial was significantly predicted by cough counts (β=.74, P=.004), explaining 76% of the variance. Particularly, coughs were associated with ED visits (r=.47. P=.004), asthma specialist visits (r=.45, P=.005), and office visits due to worsening asthma (r=.39, P=.02) that occurred in the 3-month period after the trial.

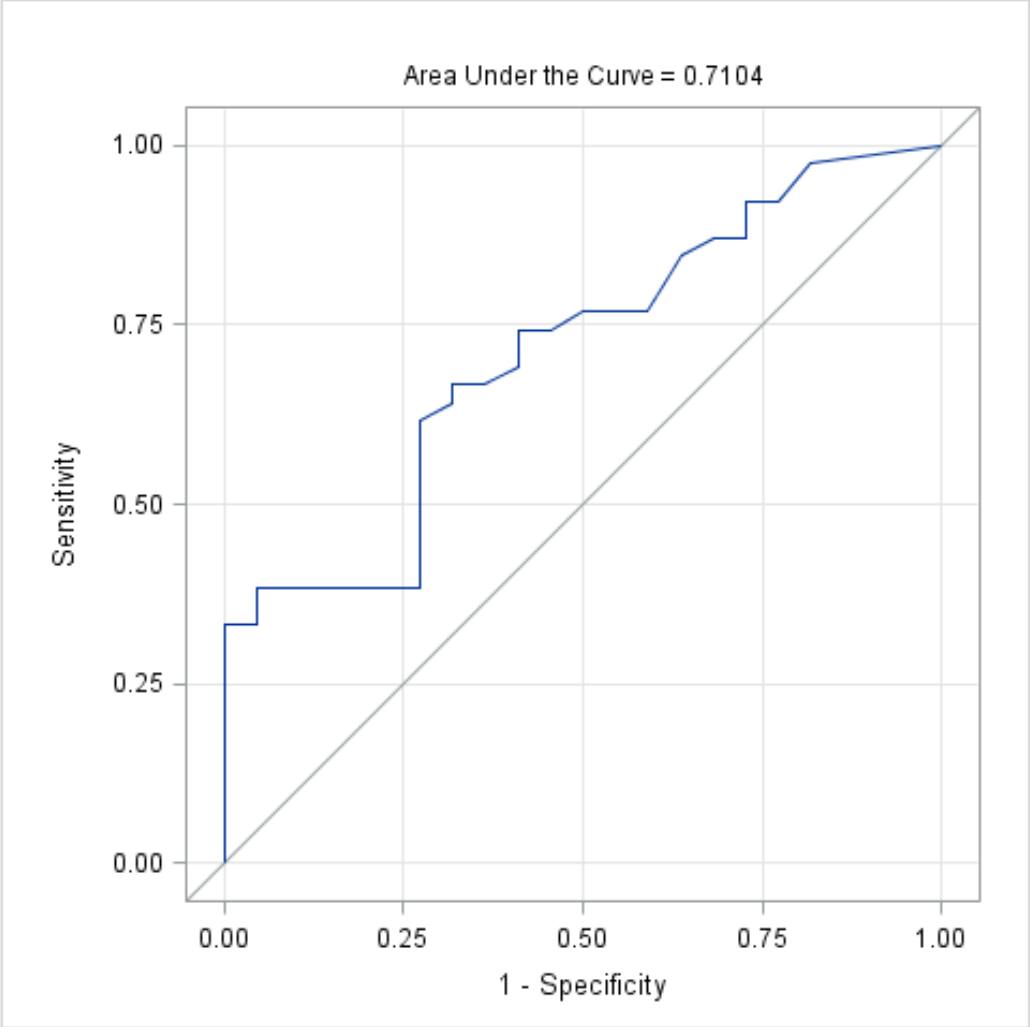
Table 3. Asthma control, quality of life, and health care utilization predicted by coughs and demographic variables.

Predictors	Asthma control		Quality of life								Health care utilization	
	β	P	Total		Activity		Symptoms		Emotional function		β	P
			β	P	β	P	β	P	β	P		
Cough counts	-.48	.003	-.55	.001	-.50	.005	-.50	.003	-.58	<.001	.74	.004
Age in years	.05	.73	.11	.47	.11	.53	.07	.65	.16	.29	.19	.42
Gender (1=female)	.08	.64	-.19	.28	-.19	.30	-.22	.20	-.11	.51	-.26	.25
Race (1=nonwhite)	.36	.14	-.03	.91	.10	.71	.04	.88	-.20	.43	.03	.90
Household income	.06	.83	.03	.92	.03	.93	.11	.68	-.09	.72	-.43	.17

Sensitivity and Specificity

The capacity of the device to distinguish participants with asthma from those without asthma was assessed using ROC curve analysis. The ROC curve for the presence of asthma diagnosis had an AUC of 0.71 (95% CI 0.58-0.84). The AUC was significantly different from chance (χ^2_1=9.7, P=.002). The optimal cutoff value of the device was 0.56, with 51.3% sensitivity and 72.7% specificity (Figure 1). The cutoff value discriminating teens with asthma from those without asthma was translated into 0.83 coughs/hour or 19.92 coughs/day.

Figure 1. Receiver operating characteristic (ROC) curve analysis for predictive values of coughs.



Overview of the Study Findings

Table 4 summarizes findings pertaining to the validity of the device. Most of the expected relationships between ADAM data and conventional measures of asthma were substantiated by our findings.

Table 4. Overview of study findings and expected relationships between cough counts and measures of asthma.

Types of validity	Expected relationships	Statistical method	Findings
Concurrent validity			
Cough counts and FeNO and lung function	Positive association with FeNO; negative association with lung function	Correlation	No significant correlations with FeNO; cough counts were negatively associated with FEV_1 and FVC
Cough counts and symptom diary data and VAS	Positive association	Correlation	Associated with limited activities and approached significance for shortness of breath and number of rescue medications use in the past 24 hours
Cough counts and asthma control	Negative association	Correlation	Cough counts were negatively associated with asthma control
Cough counts and quality of life	Negative association	Correlation	Approached significance with quality of life, activity, and symptom subscales
Cough counts and health care utilization	Positive association	Correlation	No association between cough counts and health care use before the 7-day trial; however, cough counts showed positive association with health care use during the 7-day trial
Predictive validity			
Cough counts and asthma control and quality of life 3 months later	Cough counts predicting asthma control and quality of life	Multiple regression	Coughs predicted asthma control 3 months later explaining 42% of the variance in asthma control. Coughs predicted the quality of life total score and each of subscales 3 months later, explaining variance in quality of life, which ranged from 28% to 41%
Cough counts and health care utilization 3 months later	Cough counts predicting health care utilization	Multiple regression	Coughs predicted health care utilization 3 months later explaining 76% of the variance in health care utilization
Clinical prediction			
Area under the curve		ROC curve analysis	0.71 (95% CI 0.58-0.84)
Cutoff point		ROC curve analysis	0.56 (0.83 coughs/hour or 19.92 coughs/day)
Sensitivity	Discrimination of positive asthma diagnosis by a cutoff	ROC curve analysis	51.3% sensitivity
Specificity	Discrimination of negative asthma diagnosis by a cutoff	ROC curve analysis	72.7% specificity

Discussion

This study examined the validity of ADAM, an investigational device that automatically monitors coughs continuously in adolescents with asthma. To our knowledge, ADAM is the first fully automated portable device facilitating continuous monitoring of the frequency of coughs that involves real-time processing, analysis, recording, and displaying of symptoms. Previously, we reported technical details [35] and the acceptance of the device by teen users [34]. In these earlier reports, we demonstrated the feasibility of developing an algorithm for coughs, but not for wheezes due to the wide intrapersonal and interpersonal variability of the acoustic signature of wheezing. Therefore, ADAM was evaluated solely as a cough-monitoring device at this time. Coughs are widely recognized as a key symptom of asthma [9,44] and the most common symptom of uncontrolled asthma in children and adolescents [7,45-47]. Asthma patients report coughs as the most troublesome symptom in their lives and as a symptom of greater importance [48]. Given the importance of coughing in asthma, ADAM can be a clinically useful monitoring tool not only for the symptom itself, but also for symptom burdens on individuals. In this paper, we examined the validity of the device as an asthma-monitoring tool and its capacity to discriminate asthma cases from controls.

Principal Results

Overall, we found positive temporal correlations, albeit low to moderate, between device data (ie, cough counts) and conventional measures of asthma symptoms and symptom control. Similarly, previous studies reported modest correlations between objective cough rates and subjective measures specific to cough (VAS for cough and cough scores) in individuals with asthma [49,50]. In our study, the generic self-report measures of asthma symptoms made it difficult to assess the degree of agreement between the device's cough counts and the amount of coughs perceived by individuals. Nonetheless, when conceptualizing coughs as an indicator of asthma condition, use of generic measures of asthma control to establish the device's validity as an asthma-monitoring tool can be justified. The demonstrated relationships between cough counts and concomitantly assessed activity limitation and other symptoms

as well as overall asthma control provide support for the validity of the device as a tool for overall asthma monitoring rather than simply for cough. As in another study [49], the relationship between cough counts and quality of life suggest that experiencing coughing can take a toll on quality of life in teens with asthma. Moreover, the positive relationships between cough counts and health care utilization and school absenteeism during the trial could be further evidence that cough counts as measured by ADAM is a compelling indicator of asthma morbidity.

Unlike earlier reports of no relationship between spirometry data and coughs measured by an objective method [49,51] or self-report [52], we found that the cough counts by ADAM were associated with poor pulmonary function (as indicated by FEV_1 and FVC), suggesting that the number of coughs are indicative of airway obstruction. However, consistent with an earlier study [49], we were unable to establish the relationship between coughing and airway inflammation measured by FeNO. This may be because most (>50%) of the asthma group were asymptomatic (ie, no indication of active airway inflammation) before and during our trial period. Replication of the trial using a large number of patients with active symptoms is needed to determine the nature of relationship between coughs and airway inflammation.

This study demonstrated strong evidence for predictive validity of the device. After adjusting for sociodemographic factors, cough counts were predictive of asthma morbidity and quality of life at 3 months after the 7-day trial. The cough count predicted poor asthma control and quality of life at 3-month post-trial, accounting for 42% and 38% of the variance, respectively. Health care services used during 3 months after the trial were predicted by the cough count by our device and explained as much as 76% of the variance in acute health care utilization 3 months later. In an earlier study [53], worsening coughing was found to be predictive of severe asthma after 9 years in adult patients. Our findings support not only ADAM's predictive validity, but also coughs as an important harbinger for upcoming deterioration of teens' asthma morbidity that could undermine their overall well-being and impose serious burdens on the health care system. As such, the findings underscore the importance and need for a monitoring device such as ADAM that can increase patients' awareness of the patterns of cough for early detection and has the potential for preventing serious and costly future consequences of asthma morbidity.

The device's capacity to discriminate correctly the asthma group from those without asthma was assessed to determine sensitivity and specificity. We found poor diagnostic sensitivity of ADAM; that is, the chances of correctly identifying those with asthma were only 51% using 20 coughs/day as a cutoff. The poor diagnostic sensitivity of the device may have been due to the asymptomatic state of more than 50% of the asthma group in which the average cough counts were below the cutoff making it difficult for the device to appropriately identify asthma cases. In the future, studies maximizing differences between subsamples in symptom presentation will be essential for adequately assessing the sensitivity of the device. By contrast, ADAM demonstrated relatively better specificity such that the device correctly classified those who did not have asthma for 73% of cases. Using the device in an environment replete with everyday noises may result in a high number of cough counts exceeding the cutoff, even in those without asthma; thus, incorrectly excluding them from the nonasthma group. We observed occasional false events (2 coughs/hour) depending on environmental noises (70% sensitivity of the cough algorithm) [35], which is strikingly similar to 2.5 events/hour by the Leicester Cough Monitor [54]. Given the poor capacity to correctly classify patients with asthma, ADAM is *not* suitable for determining a clinical diagnosis of asthma, but is intended solely for monitoring cough in those with a confirmed asthma diagnosis. Like any monitoring device, false alarms are still an issue because these can cause unnecessary concerns or undermine users' confidence in the monitoring tool. Therefore, further optimizing the accuracy of the monitoring device is warranted by refining the algorithm and adopting noninvasive techniques to minimize the influence of environmental noises (eg, direct application of an adhesive microphone on the chest wall).

Study Limitations

Several limitations to this study's design warrant caution. Because the device stored only the number of coughs without sound recording, validating the accuracy of cough counts through the manual confirmation of corresponding cough events was not done. Except for pulmonary function tests and FeNO, we primarily used self-report measures that were not specific to coughs to establish the validity of the device. The generic nature of the measure may have contributed to some of the nonsignificant correlations between the measures and device data. Although self-report measures are inherently subject to recollection bias, we attempted to address the challenge by strategically collecting daily symptom data electronically. In health care utilization data, we observed more than 90% agreement between self-report and medical record review. Therefore, it appears that recollection bias played little influence on our validity outcomes. Nonetheless, future research is needed to evaluate the validity of the device by simultaneously recording raw cough sounds and by using cough specific measures.

Moreover, distinction between the asthma and nonasthma groups was blurred because most of the asthma group did not present active symptoms during the trial, which may account for the device's poor sensitivity. Also, use of the small and convenient sample of adolescents limited generalizability of the findings. In addition, the length of the 7-day trial might not be long enough to observe any meaningful changes in asthma symptoms, health care utilization, or biological measures including FEV_1 and FeNO. A longer observation period is warranted to assess the extent to which the device adequately captures changes over time in symptom patterns and other measures of asthma status. Lastly, there were a few technical challenges and limitations to the optimal operation of the device, which we discussed in our earlier report [34].

Comparison With Prior Work

Several approaches have been attempted to develop technologies to monitor symptoms objectively, particularly coughs [45,49,50,54-58]. However, these existing approaches are

considered unsuitable for ambulatory monitoring of symptoms due to practical challenges, including the laborious and time-intensive nature of processing audio data (so not fully automated and unable to provide real-time feedback to users), inability to monitor continuously beyond 24 hours, and the conspicuous appearance of the systems. For instance, the Leicester Cough Monitor [54,59,60] uses a similar sound recognition technical approach to that of ADAM. Unlike ADAM, in the Leicester Cough Monitor, the ambulatory component of the system consists strictly of an audio recorder and audio analysis is performed offline, which can take approximately 1 hour to process a 24-hour audio recording. This differs fundamentally from ADAM in which all of the processing/annotation of audio data are performed instantly on the mobile device and the results (cough counts) are provided as feedback in real time. Other cough detection systems [61] are not intended for ambulatory use or have yet to be validated in vivo for an extended observation period. In that sense, ADAM is the first ambulatory cough-monitoring device on a consumer mobile system with a capacity to fully automatize continuous real-time processing. This is also the first attempt to evaluate the validity of a cough detection device as a symptom-monitoring tool for several consecutive days in adolescents with asthma. Using a popular mobile device, such as iPod, as a platform for ADAM was well received by adolescents with asthma [34] and the majority of participants used the device daily for a week or longer period supporting ADAM as a sustainable asthma-monitoring tool for adolescents.

Conclusions

Overall, this study demonstrated the validity of ADAM as a symptom-monitoring device in teens with a confirmed asthma diagnosis. Poorly controlled asthma takes a toll on teens' overall health and quality of life as well as on the health care system due to an increased economic burden associated with the use of urgent types of health care services. ADAM can potentially mitigate the adverse consequences by helping users detect and treat early symptoms before advancing to a poorly controlled state. The device is useful in increasing understanding about one's current status of asthma control and in predicting asthma morbidity and quality of life for the near future. Such information can make the users become aware of symptoms and triggers and enable them to take appropriate and timely actions to address symptoms or prevent further deterioration of their symptoms. Objective symptom information from the device would be clinically useful in establishing optimum treatment plans and evaluating treatment effects. ADAM can be particularly useful and effective in monitoring symptoms occurring at night when environmental noises are minimal. Nighttime coughs were more common than wheezing [62] and worsening of asthma symptoms often occurs during nighttime [63]. As such, nighttime symptoms are often indicative of poorly controlled asthma [31,64], but patients often do not recognize or tend to neglect to report nighttime symptoms [65-68]. ADAM can be an invaluable tool that monitors nighttime symptoms, which would provide important clinical insight into the degree of symptom control and the response to asthma treatment.

Nonetheless, we caution that the device should not be considered as a diagnostic tool and its application to a broader age range remains to be evaluated. Although we suggest coughing as an important indicator of asthma morbidity, understanding the symptom in the context of other information, such as activity level and medication use, can potentially augment its clinical value and relevance. Originally, ADAM was designed to monitor symptoms and activity levels simultaneously capitalizing on the host system's (iPod) built-in accelerometer to offer an insight into the nature of the relationships between symptoms and activity levels (eg, exercise-induced asthma). However, continuously recording and processing sounds and activity simultaneously quickly drained the platform battery, posing challenges for long-term uninterrupted sound processing. More study is needed to reassess the feasibility of concurrent operation of these 2 applications with optimized code and application power management. ADAM also allowed users to record the use of medications throughout the day to help them systemically review changes in symptom patterns in relation to medication use. This can motivate users' adherence to treatment plans (when they noted symptom reduction after medication) or inform users or providers of the need for adjusting current medication or dosage (when no relief from symptom was achieved after medication). A clinical trial is needed to examine the effects of the device's medication tracking function on users' medication adherence.

ADAM represents an important undertaking in the field of mobile health (mHealth) that has exploded in the past decade to improve health outcomes and patients' self-management capacity [69-71]. The application of mobile technologies in daily assessment of asthma symptoms can be a particularly attractive option for adolescents. Yet to be determined is the extent to which the appeal of mobile technology can translate into better symptom monitoring, awareness, and self-management behaviors, which ultimately leads to improved asthma outcomes.

Potential Clinical Application of ADAM: Sample Case Scenario

The following sample clinical scenario demonstrates the potential application of ADAM in clinical practice:

> *JD is a 16-year-old boy with moderate persistent asthma and a history of several emergency visits for asthma each year. His provider had prescribed a daily controller medication for him, but he rarely uses it because he has not perceived any significant benefit. He reports not sleeping well, which causes him to feel poorly rested in the morning. He relies heavily on albuterol with several self-administered doses each day. After a discussion with his provider, JD agreed to use ADAM for a month. At a follow-up visit, the provider and JD reviewed the symptom pattern displayed on the screen of the mobile phone monitor. Data from the device revealed significant coughs, predominantly at night, which JD was unaware. After discussing the results, JD agreed to use the controller medication daily to see if the symptoms improved. After 1 month, JD repeated his symptom monitoring with ADAM and discovered his coughing events were significantly decreased compared to the prior*

monitoring period. In addition, JD was alerted by ADAM when his symptoms became out of control, which allowed him to take precautions and closely follow his asthma action plan. He agreed to continue to use the controller medication for another 2 months and to use ADAM intermittently to help track his symptoms.

Acknowledgments

This study was funded by the National Institute of Health/National Institute for Nursing Research (R01NR011169-01A1).

Authors' Contributions

HR, as the principal investigator, was responsible for the overall study and drafted the entire manuscript; MJB analyzed the data and contributed to the "Results" section; MS participated in technology development and provided input in describing technical aspects of the device. MFB led and oversaw the technology development for ADAM and contributed to the critical review of the manuscript. We thank Ms Eileen Fairbanks for having coordinated and managed the study on a daily basis and K Comeau, S Miner, and J Mammen for having contributed to participant recruitment and data collection.

Conflicts of Interest

None declared.

References

1. Bloom B, Cohen RA, Freeman G. Summary health statistics for US children: National Health Interview Survey, 2010. Vital Health Stat 10 2011 Dec(250):1-80 [FREE Full text] [Medline: 22338334]
2. Akinbami LJ, Moorman JE, Garbe PL, Sondik EJ. Status of childhood asthma in the United States, 1980-2007. Pediatrics 2009 Mar;123 Suppl 3:S131-S145. [doi: 10.1542/peds.2008-2233C] [Medline: 19221156]
3. Gibson PG, Ram FS, Powell H. Asthma education. Respir Med 2003 Sep;97(9):1036-1044. [Medline: 14509558]
4. Warsi A, Wang PS, LaValley MP, Avorn J, Solomon DH. Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. Arch Intern Med 2004;164(15):1641-1649. [doi: 10.1001/archinte.164.15.1641] [Medline: 15302634]
5. Bruzzese J, Stepney C, Fiorino EK, Bornstein L, Wang J, Petkova E, et al. Asthma self-management is sub-optimal in urban Hispanic and African American/black early adolescents with uncontrolled persistent asthma. J Asthma 2012 Feb;49(1):90-97. [doi: 10.3109/02770903.2011.637595] [Medline: 22149141]
6. Bruzzese J, Bonner S, Vincent EJ, Sheares BJ, Mellins RB, Levison MJ, et al. Asthma education: the adolescent experience. Patient Educ Couns 2004 Dec;55(3):396-406. [doi: 10.1016/j.pec.2003.04.009] [Medline: 15582346]
7. Davis KJ, Disantostefano R, Peden DB. Is Johnny wheezing? Parent-child agreement in the Childhood Asthma in America survey. Pediatr Allergy Immunol 2011 Feb;22(1 Pt 1):31-35. [doi: 10.1111/j.1399-3038.2010.01016.x] [Medline: 21261742]
8. Bhogal S, Zemek R, Ducharme FM. Written action plans for asthma in children. Cochrane Database Syst Rev 2006(3):CD005306. [doi: 10.1002/14651858.CD005306.pub2] [Medline: 16856090]
9. National Asthma Education and Prevention Program. Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007. J Allergy Clin Immunol 2007 Nov;120(5 Suppl):S94-138. [doi: 10.1016/j.jaci.2007.09.043] [Medline: 17983880]
10. Horak E, Grässl G, Skladal D, Ulmer H. Lung function and symptom perception in children with asthma and their parents. Pediatr Pulmonol 2003 Jan;35(1):23-28. [doi: 10.1002/ppul.10218] [Medline: 12461735]
11. Fritz GK, McQuaid EL, Spirito A, Klein RB. Symptom perception in pediatric asthma: relationship to functional morbidity and psychological factors. J Am Acad Child Adolesc Psychiatry 1996 Aug;35(8):1033-1041. [doi: 10.1097/00004583-199608000-00014] [Medline: 8755800]
12. Cabral AL, Conceição GM, Saldiva PH, Martins MA. Effect of asthma severity on symptom perception in childhood asthma. Braz J Med Biol Res 2002 Mar;35(3):319-327 [FREE Full text] [Medline: 11887209]
13. Baker RR, Mishoe SC, Zaitoun FH, Arant CB, Lucas J, Rupp NT. Poor perception of airway obstruction in children with asthma. J Asthma 2000;37(7):613-624. [Medline: 11059529]
14. Mittal V, Khanna P, Panjabi C, Shah A. Subjective symptom perceptual accuracy in asthmatic children and their parents in India. Ann Allergy Asthma Immunol 2006 Oct;97(4):484-489. [doi: 10.1016/S1081-1206(10)60939-8] [Medline: 17069103]
15. Rietveld S, Everaerd W. Perceptions of asthma by adolescents at home. Chest 2000 Feb;117(2):434-439. [Medline: 10669687]
16. Halterman JS, Yoos HL, Kitzman H, Anson E, Sidora-Arcoleo K, McMullen A. Symptom reporting in childhood asthma: a comparison of assessment methods. Arch Dis Child 2006 Sep;91(9):766-770 [FREE Full text] [doi: 10.1136/adc.2006.096123] [Medline: 16705016]
17. Fuhlbrigge AL, Guilbert T, Spahn J, Peden D, Davis K. The influence of variation in type and pattern of symptoms on assessment in pediatric asthma. Pediatrics 2006 Aug;118(2):619-625. [doi: 10.1542/peds.2005-2963] [Medline: 16882815]

18. Rhee H, Belyea MJ, Elward KS. Patterns of asthma control perception in adolescents: associations with psychosocial functioning. J Asthma 2008 Sep;45(7):600-606 [FREE Full text] [doi: 10.1080/02770900802126974] [Medline: 18773334]
19. Britto MT, Byczkowski TL, Hesse EA, Munafo JK, Vockell AB, Yi MS. Overestimation of impairment-related asthma control by adolescents. J Pediatr 2011 Jun;158(6):1028-1030.e1. [doi: 10.1016/j.jpeds.2011.01.034] [Medline: 21392791]
20. Arga M, Sahbaz H, Bakirtas A, Turktas I, Demirsoy MS. Does self-monitoring by means of symptom diaries improve asthma control in children? J Asthma 2014 Apr;51(3):299-305. [doi: 10.3109/02770903.2013.867974] [Medline: 24274828]
21. Bijl-Hofland ID, Cloosterman SG, Folgering HT, Akkermans RP, van Schayck CP. Relation of the perception of airway obstruction to the severity of asthma. Thorax 1999 Jan;54(1):15-19 [FREE Full text] [Medline: 10343625]
22. Koh YI, Choi IS, Lim H. Airway responsiveness as a direct factor contributing to the dyspnoea perception in asthma. Respir Med 2001 Jun;95(6):464-470 [FREE Full text] [doi: 10.1053/rmed.2001.1068] [Medline: 11421503]
23. Veen JC, Smits HH, Ravensberg AJ, Hiemstra PS, Sterk PJ, Bel EH. Impaired perception of dyspnea in patients with severe asthma. Relation to sputum eosinophils. Am J Respir Crit Care Med 1998 Oct;158(4):1134-1141. [doi: 10.1164/ajrccm.158.4.9710087] [Medline: 9769272]
24. Brouwer AF, Roorda RJ, Brand PL. Home spirometry and asthma severity in children. Eur Respir J 2006 Dec;28(6):1131-1137 [FREE Full text] [doi: 10.1183/09031936.06.00118205] [Medline: 16870659]
25. Slieker MG, van der Ent CK. The diagnostic and screening capacities of peak expiratory flow measurements in the assessment of airway obstruction and bronchodilator response in children with asthma. Monaldi Arch Chest Dis 2003;59(2):155-159. [Medline: 14635506]
26. Nazir Z, Razaq S, Mir S, Anwar M, Al MG, Sajad M, et al. Revisiting the accuracy of peak flow meters: a double-blind study using formal methods of agreement. Respir Med 2005 May;99(5):592-595 [FREE Full text] [doi: 10.1016/j.rmed.2004.10.015] [Medline: 15823456]
27. Kamps AW, Roorda RJ, Brand PL. Peak flow diaries in childhood asthma are unreliable. Thorax 2001 Mar;56(3):180-182 [FREE Full text] [Medline: 11182008]
28. Redline S, Wright EC, Kattan M, Kercsmar C, Weiss K. Short-term compliance with peak flow monitoring: results from a study of inner city children with asthma. Pediatr Pulmonol 1996 Apr;21(4):203-210. [doi: 10.1002/(SICI)1099-0496(199604)21:4<203::AID-PPUL1>3.0.CO;2-P] [Medline: 9121848]
29. Yoos HL, Kitzman H, McMullen A, Henderson C, Sidora K. Symptom monitoring in childhood asthma: a randomized clinical trial comparing peak expiratory flow rate with symptom monitoring. Ann Allergy Asthma Immunol 2002 Mar;88(3):283-291. [doi: 10.1016/S1081-1206(10)62010-8] [Medline: 11926622]
30. Burkhart PV, Dunbar-Jacob JM, Rohay JM. Accuracy of children's self-reported adherence to treatment. J Nurs Scholarsh 2001;33(1):27-32. [Medline: 11253577]
31. van der Meer V, Rikkers-Mutsaerts E R V ER, Sterk PJ, Thiadens HA, Assendelft WJ, Sont JK. Compliance and reliability of electronic PEF monitoring in adolescents with asthma. Thorax 2006 May;61(5):457-458 [FREE Full text] [doi: 10.1136/thx.2005.054619] [Medline: 16648355]
32. Wensley D, Silverman M. Peak flow monitoring for guided self-management in childhood asthma: a randomized controlled trial. Am J Respir Crit Care Med 2004 Sep 15;170(6):606-612. [doi: 10.1164/rccm.200307-1025OC] [Medline: 15184205]
33. Adams RJ, Boath K, Homan S, Campbell DA, Ruffin RE. A randomized trial of peak-flow and symptom-based action plans in adults with moderate-to-severe asthma. Respirology 2001 Dec;6(4):297-304. [Medline: 11844120]
34. Rhee H, Miner S, Sterling M, Halterman JS, Fairbanks E. The development of an automated device for asthma monitoring for adolescents: methodologic approach and user acceptability. JMIR Mhealth Uhealth 2014;2(2):e27 [FREE Full text] [doi: 10.2196/mhealth.3118] [Medline: 25100184]
35. Sterling M, Rhee H, Bocko M. Automated cough assessment on a mobile platform. J Med Eng 2014;2014:1-9 [FREE Full text] [doi: 10.1155/2014/951621] [Medline: 25506590]
36. Smith AD, Taylor DR. Is exhaled nitric oxide measurement a useful clinical test in asthma? Curr Opin Allergy Clin Immunol 2005 Feb;5(1):49-56. [Medline: 15643344]
37. Harkins MS, Fiato K, Iwamoto GK. Exhaled nitric oxide predicts asthma exacerbation. J Asthma 2004 Jun;41(4):471-476. [Medline: 15281333]
38. Nordvall SL, Janson C, Kalm-Stephens P, Foucard T, Torén K, Alving K. Exhaled nitric oxide in a population-based study of asthma and allergy in schoolchildren. Allergy 2005 Apr;60(4):469-475. [doi: 10.1111/j.1398-9995.2005.00735.x] [Medline: 15727578]
39. Spergel JM, Fogg MI, Bokszczanin-Knosala A. Correlation of exhaled nitric oxide, spirometry and asthma symptoms. J Asthma 2005 Dec;42(10):879-883. [doi: 10.1080/02770900500371344] [Medline: 16393728]
40. American Thoracic Society, European European Respiratory Society. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. Am J Respir Crit Care Med 2005 Apr 15;171(8):912-930. [doi: 10.1164/rccm.200406-710ST] [Medline: 15817806]
41. American Thoracic Society. Standardization of spirometry, 1994 Update. Am J Respir Crit Care Med 1995 Sep;152(3):1107-1136. [doi: 10.1164/ajrccm.152.3.7663792] [Medline: 7663792]
42. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in children with asthma. Qual Life Res 1996 Feb;5(1):35-46. [Medline: 8901365]

43. Juniper EF, O'Byrne PM, Ferrie PJ, King DR, Roberts JN. Measuring asthma control. Clinic questionnaire or daily diary? Am J Respir Crit Care Med 2000 Oct;162(4 Pt 1):1330-1334. [doi: 10.1164/ajrccm.162.4.9912138] [Medline: 11029340]
44. Janson C, Chinn S, Jarvis D, Burney P. Determinants of cough in young adults participating in the European Community Respiratory Health Survey. Eur Respir J 2001 Oct;18(4):647-654 [FREE Full text] [Medline: 11716169]
45. Rietveld S, Rijssenbeek-Nouwens LH. Diagnostics of spontaneous cough in childhood asthma: results of continuous tracheal sound recording in the homes of children. Chest 1998 Jan;113(1):50-54. [Medline: 9440567]
46. Wildhaber J, Carroll WD, Brand PL. Global impact of asthma on children and adolescents' daily lives: the room to breathe survey. Pediatr Pulmonol 2012 Apr;47(4):346-357. [doi: 10.1002/ppul.21557] [Medline: 22028276]
47. Rhee H, Fairbanks E, Butz A. Symptoms, feelings, activities and medication use in adolescents with uncontrolled asthma: lessons learned from asthma diaries. J Pediatr Nurs 2014;29(1):39-46 [FREE Full text] [doi: 10.1016/j.pedn.2013.04.009] [Medline: 23685266]
48. Osman LM, McKenzie L, Cairns J, Friend JA, Godden DJ, Legge JS, et al. Patient weighting of importance of asthma symptoms. Thorax 2001 Feb;56(2):138-142 [FREE Full text] [Medline: 11209103]
49. Marsden PA, Smith JA, Kelsall AA, Owen E, Naylor JR, Webster D, et al. A comparison of objective and subjective measures of cough in asthma. J Allergy Clin Immunol 2008 Nov;122(5):903-907. [doi: 10.1016/j.jaci.2008.08.029] [Medline: 18842290]
50. Hsu JY, Stone RA, Logan-Sinclair RB, Worsdell M, Busst CM, Chung KF. Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder. Eur Respir J 1994 Jul;7(7):1246-1253 [FREE Full text] [Medline: 7925902]
51. Hoskyns EW, Beardsmore CS, Simpson H. Chronic night cough and asthma severity in children with stable asthma. Eur J Pediatr 1995 Apr;154(4):320-325. [Medline: 7607286]
52. Teeter JG, Bleecker ER. Relationship between airway obstruction and respiratory symptoms in adult asthmatics. Chest 1998 Feb;113(2):272-277. [Medline: 9498938]
53. de Marco R, Marcon A, Jarvis D, Accordini S, Almar E, Bugiani M, European Community Respiratory Health Survey Therapy Group. Prognostic factors of asthma severity: a 9-year international prospective cohort study. J Allergy Clin Immunol 2006 Jun;117(6):1249-1256. [doi: 10.1016/j.jaci.2006.03.019] [Medline: 16750983]
54. Birring SS, Fleming T, Matos S, Raj AA, Evans DH, Pavord ID. The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough. Eur Respir J 2008 May;31(5):1013-1018 [FREE Full text] [doi: 10.1183/09031936.00057407] [Medline: 18184683]
55. Munyard P, Busst C, Logan-Sinclair R, Bush A. A new device for ambulatory cough recording. Pediatr Pulmonol 1994 Sep;18(3):178-186. [Medline: 7800435]
56. Hamutcu R, Francis J, Karakoc F, Bush A. Objective monitoring of cough in children with cystic fibrosis. Pediatr Pulmonol 2002 Nov;34(5):331-335. [doi: 10.1002/ppul.10174] [Medline: 12357476]
57. Smith J, Woodcock A. New developments in the objective assessment of cough. Lung 2008;186 Suppl 1:S48-S54. [doi: 10.1007/s00408-007-9059-1] [Medline: 18066694]
58. Amoh J, Odame K. Technologies for developing ambulatory cough monitoring devices. Crit Rev Biomed Eng 2013;41(6):457-468. [Medline: 24940659]
59. Matos S, Birring SS, Pavord ID, Evans DH. Detection of cough signals in continuous audio recordings using hidden Markov models. IEEE Trans Biomed Eng 2006 Jun;53(6):1078-1083. [doi: 10.1109/TBME.2006.873548] [Medline: 16761835]
60. Matos S, Birring SS, Pavord ID, Evans DH. An automated system for 24-h monitoring of cough frequency: the leicester cough monitor. IEEE Trans Biomed Eng 2007 Aug;54(8):1472-1479. [doi: 10.1109/TBME.2007.900811] [Medline: 17694868]
61. Shin S, Hashimoto T, Hatano S. Automatic detection system for cough sounds as a symptom of abnormal health condition. IEEE Trans Inf Technol Biomed 2009 Jul;13(4):486-493. [doi: 10.1109/TITB.2008.923771] [Medline: 19273017]
62. Manfreda J, Becklake MR, Sears MR, Chan-Yeung M, Dimich-Ward H, Siersted HC, et al. Prevalence of asthma symptoms among adults aged 20-44 years in Canada. CMAJ 2001 Apr 3;164(7):995-1001 [FREE Full text] [Medline: 11314453]
63. Ballard RD, Saathoff MC, Patel DK, Kelly PL, Martin RJ. Effect of sleep on nocturnal bronchoconstriction and ventilatory patterns in asthmatics. J Appl Physiol (1985) 1989 Jul;67(1):243-249. [Medline: 2759949]
64. Bentur L, Beck R, Shinawi M, Naveh T, Gavriely N. Wheeze monitoring in children for assessment of nocturnal asthma and response to therapy. Eur Respir J 2003 Apr;21(4):621-626 [FREE Full text] [Medline: 12762346]
65. Meijer GG, Postma DS, Wempe JB, Gerritsen J, Knol K, van Aalderen WM. Frequency of nocturnal symptoms in asthmatic children attending a hospital out-patient clinic. Eur Respir J 1995 Dec;8(12):2076-2080 [FREE Full text] [Medline: 8666103]
66. Storms WW, Bodman SF, Nathan RA, Byer P. Nocturnal asthma symptoms may be more prevalent than we think. J Asthma 1994;31(4):313-318. [Medline: 8040155]
67. Baughman RP, Loudon RG. Lung sound analysis for continuous evaluation of airflow obstruction in asthma. Chest 1985 Sep;88(3):364-368. [Medline: 4028846]
68. Raherison C, Abouelfath A, Le Gros V, Taytard A, Molimard M. Underdiagnosis of nocturnal symptoms in asthma in general practice. J Asthma 2006 Apr;43(3):199-202. [doi: 10.1080/02770900600566744] [Medline: 16754521]

69. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
70. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
71. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]

Abbreviations

ADAM: Automated Device for Asthma Monitoring
ATS: American Thoracic Society
AUC: area under the curve
FeNO: fractional exhaled nitric oxide
FEV_1: forced expiratory volume in first second of expiration
FVC: forced vital capacity
ICS: inhaled corticosteroids
NHLBI: National Heart, Lung, and Blood Institute
PAQLQ: Pediatric Asthma Quality of Life Questionnaire
ROC: receiver operating characteristic
SABA: short-acting beta agonists
VAS: Visual Analog Scale

Edited by G Eysenbach; submitted 23.07.15; peer-reviewed by N Diviani, P Schulz; comments to author 05.08.15; revised version received 18.08.15; accepted 19.08.15; published 16.10.15.

Please cite as:
Rhee H, Belyea MJ, Sterling M, Bocko MF
Evaluating the Validity of an Automated Device for Asthma Monitoring for Adolescents: Correlational Design
J Med Internet Res 2015;17(10):e234
URL: http://www.jmir.org/2015/10/e234/
doi:10.2196/jmir.4975
PMID:26475634

Viewpoint

The Potential Role of Social Media Platforms in Community Awareness of Antibiotic Use in the Gulf Cooperation Council States: Luxury or Necessity?

Hosam Mamoon Zowawi[1,2,3], BSc, MSc (Hons); Malak Abedalthagafi[4], MD, MBA; Florie A Mar[5], PhD; Turki Almalki[6], MBBS; Abdullah H Kutbi[6]; Tiffany Harris-Brown[1], RN, MPH; Stephan Harbarth[7], MD, MPH; Hanan H Balkhy[2,3], MD; David L Paterson[1], MBBS, PhD; Rihab Abdalazez Hasanain[8], BSc, MA (Hons)

[1]The University of Queensland, UQ Centre for Clinical Research, Herston QLD 4029, Australia

[2]King Saud bin Abdulaziz University for Health Sciences, College of Medicine, Riyadh, Saudi Arabia

[3]WHO Collaborating Centre for Infection Prevention and Control and GCC Center for Infection Control, Riyadh, Saudi Arabia

[4]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, MA, United States

[5]Department of Pathology, University of California, San Francisco, CA 94143, USA, CA, United States

[6]Faculty of Medicine, Taif University, Taif, Saudi Arabia

[7]Infection Control Programme and WHO Collaborating Centre on Patient Safety, University Hospitals of Geneva, Rue Gabrielle-Perret-Gentil 4,, Switzerland

[8]School of Public Health, Queensland University of Technology, QLD, Australia

Corresponding Author:
Rihab Abdalazez Hasanain, BSc, MA (Hons)
School of Public Health
Queensland University of Technology
Victoria Park Rd Kelvin Grove QLD 4059
QLD,
Australia
Phone: 61 7 3138 5879
Fax: 61 7 3138 5880
Email: Rihab.Hasanain@student.qut.edu.au

Abstract

The increasing emergence and spread of antimicrobial resistance (AMR) is a serious public health issue. Increasing the awareness of the general public about appropriate antibiotic use is a key factor for combating this issue. Several public media campaigns worldwide have been launched; however, such campaigns can be costly and the outcomes are variable and difficult to assess. Social media platforms, including Twitter, Facebook, and YouTube, are now frequently utilized to address health-related issues. In many geographical locations, such as the countries of the Gulf Cooperation Council (GCC) States (Saudi Arabia, United Arab Emirates, Kuwait, Oman, Qatar, and Bahrain), these platforms are becoming increasingly popular. The socioeconomic status of the GCC states and their reliable communication and networking infrastructure has allowed the penetration and scalability of these platforms in the region. This might explain why the Saudi Ministry of Health is using social media platforms alongside various other media platforms in a large-scale public awareness campaign to educate at-risk communities about the recently emerged Middle East respiratory syndrome coronavirus (MERS-CoV). This paper discusses the potential for using social media tools as cost-efficient and mass education platforms to raise awareness of appropriate antibiotic use in the general public and in the medical communities of the Arabian Peninsula.

(J Med Internet Res 2015;17(10):e233) doi:10.2196/jmir.3891

KEYWORDS

social media; antibiotics; awareness; health campaigns; Gulf States

Introduction

Antibiotic resistance has become a severe public health threat worldwide, including in the Gulf Cooperation Council (GCC) states [1]. Hence, various initiatives across the globe have been launched to combat this issue. In 2011, the World Health Organization (WHO) themed its annual day to address this issue with the slogan "No action today, no cure tomorrow" and listed the actions to be undertaken. These actions included providing education to achieve effective antibiotic use [2]. Antibiotic misuse, such as demanding unneeded antibiotics, purchasing antibiotics over-the-counter without a prescription, and not completing a course of antibiotics, is associated with the emergence and selection of antibiotic-resistant bacteria [3]. Nosocomial infections caused by antibiotic-resistant pathogens are significantly associated with an increased length of hospital stay and increased cost [4].

European surveys have shown that the general public still misunderstands the function and correct use of antibiotics [5,6]. On many occasions, the WHO has highlighted the importance of involving the general public alongside health care professionals for combating the emergence of antimicrobial resistance (AMR) [7,8]. To alleviate the public's lack of cognizance about AMR, various countries in Europe [7,9], as well as the United States [10] and other countries, have initiated public campaigns to raise awareness about the appropriate use of antibiotics.

Considering Saudi Arabia and other Gulf countries, various studies have addressed the issue of antibiotic misuse in hospital settings and the easy over-the-counter access to prescription antibiotics. For example, antibiotic use in intensive care units (ICUs) in Saudi Arabia has been found to be 10 times greater than that in the United States and some European countries [1,11-13]. This issue also extends to the wider GCC community. A Saudi study found that 77.6% of pharmacies dispensed antibiotics without a prescription primarily to treat scenarios consistent with viral infections [14], whereas 68.4% of antibiotics from Abu Dhabi pharmacies were sold over-the-counter [15], suggesting lack of antibiotics knowledge [16].

These examples of antibiotics misuse and others suggest the urgent need for a public campaign in Saudi Arabia and beyond to provide greater education on the proper use of antibiotics along with the concept of reserving antibiotics for use only when they are truly needed. In April 2015, the GCC Center for Infection Control released the multilevel GCC Strategic Plan for Combating Antimicrobial Resistance, which sets the framework for the regional and national plans [17]. One of the 5 strategic roadmaps addressed is the importance of preserving and restricting the available antimicrobial agents for human use. Interventional methods include educating antibiotic prescribers, patients, and the general public on the importance of appropriate antimicrobials use and basic infection prevention and control (eg, immunization and hygiene) [17].

Social media platforms are being widely used for health promotion advocates and to endorse traditional awareness campaigns. They have unique characteristics for sharing open access information, providing a platform for dynamic conversations with communities and social groups, and keeping users connected with their topics of interest [18]. They have been used to raise awareness for obesity, diabetes [19], and adolescent dating violence [20]. As an outreach effort, the European Antibiotic Awareness Day released a toolkit to advise on how to engage in social media activities promoting prudent antibiotic use [18]. Information provided in the toolkit relates to European countries and may not fully apply to the GCC states.

In this paper, we discuss the planning, setup, and potential effectiveness of developing a mass education campaign via social media platforms to raise general public and medical awareness of appropriate antibiotic usage in the GCC states.

The Influence of Educational Campaigns on Antibiotic Use and Antibiotic Resistance

Before initiating an educational media campaign, it is important to review the effectiveness of previous initiatives that were developed and launched in other parts of the world to raise awareness and provide guidance on responsible antibiotic use (Table 1).

Table 1. Summary of selected antibiotic awareness campaigns worldwide.

Name	Site (country)	Duration	Method used	Target audience	Reduction rate
Belgian Antibiotic Policy Coordinating Committee (BAPCOC) establishment [21]	Belgium	Launched in 1999-present	Multimedia campaigns, national campaigns, publication of clinical practice guidelines, support for the establishment of antibiotic management teams (AMTs)	Belgian community	36% reduction in outpatient antibiotic use per 1000 inhabitants per day during winter season
"Les antibiotiques c'est pas automatique" ("Antibiotics are not automatic"); part of the national campaign "Keep antibiotics working" [22]	All 22 regions across France	2001-2007	Mass media campaigns, one-on-one physician education sessions	General public and health care professionals	26.5% reduction of antibiotic prescriptions per 100 inhabitants during winter season over a 5-year period
Antibiotics Awareness Week [23]	Australia	2012-present	Facebook, Twitter, online pledging	All Australians	Unknown as yet
Local low-cost information campaign [9]	Emilia-Romagna region (Northern Italy)	November 2011-Februrary 2012	Brochures, posters, local media advertisements, and visual aids	General public	4.3% reduction in defined daily doses of prescribed antibiotics in intervention group
English public antibiotic campaigns [7]	England and Scotland	2008	Posters displayed in magazines and newspapers	General public	No improvement observed in postsurvey of public's understanding
e-Bug [24]	European countries and Saudi Arabia	Launched in 2006-present	Website-based games	Junior and senior school students	Not assessed

Data from Belgium and France have also revealed a reduction in the misuse of antibiotics after educational interventions. A decrease of 26.5% in antibiotic prescriptions was observed in France between 2002 and 2007 compared with the preintervention period (2000-2002), with the largest reduction observed in children [22]. A 36% reduction in antibiotic prescriptions was also observed in Belgium from 1999-2000 to 2006-2007 [21]. Both countries also reported a decrease in the incidence of infection with invasive penicillin- and macrolide-resistant *Streptococcus pneumoniae*. It was noted that the decrease in Belgium occurred before the wider use of the pneumococcal conjugate vaccine (PCV7), indicating that the vaccine did not contribute to this initial reduction in the incidence of infection with invasive penicillin-resistant pneumococci [5,6]. These data suggest the effectiveness of antibiotic awareness media campaigns in decreasing the use of antibiotics and hence in reducing the impact of antibiotic resistance.

Recently, Formoso et al [9] reported the effectiveness of a low-cost media campaign on antibiotic use in an Italian province that lasted for 5 months during the cold season. The intervention materials included visual aids, such as posters, brochures, and advertisements, which were used in the local media. The key messages of the campaign were codesigned by a physician practicing in the intervention area. Antibiotic prescriptions were significantly reduced by 4.3% in the intervention area compared with the control area. However, the general population's knowledge and attitudes about antibiotic resistance were not changed by the campaign [9].

In 2008, antibiotic awareness campaigns were carried out in England and Scotland by broadcasting key messages in advertising published in magazines and newspapers. Unfortunately, the campaign did not show any positive effects in either England or Scotland. No improvement was observed in the general public's understanding of antibiotic misuse to treat coughs and colds despite the fact that in 2009 more public respondents agreed that "resistance to antibiotics is a problem in British hospitals" than in 2008 [7]. In fact, this is not the only documented failure of antibiotic awareness campaigns. Huttner et al [5] reviewed 22 campaigns launched in high-income countries between 1990 and 2007. At least 3 of these campaigns failed and the effect of 3 others is unknown because of a lack of follow-up assessment on antibiotic use [5].

The e-Bug project is an example of an innovative approach to raise awareness about microbes and infection prevention. The aim of the e-Bug project is to disseminate educational materials about microbes (both beneficial and pathogenic) to junior and senior school students across Europe. The project relies on website-based gaming and entertainment-based lessons [24]. The e-Bug project now has partners in 26 different European countries, as well as Saudi Arabia, providing educational materials in different European languages and in Arabic [25]. On May 2015, the e-Bug website had a total of 17,391 visitors and the Saudi Arabian site had 76 visitors in total [26]. The impact of the e-Bug project is not clear because no evaluation of its implementation and impact on behavioral changes in targeted groups was ever performed [24].

Various factors have been suggested to be necessary to achieve success in antibiotic awareness campaigns. These include carefully designing key messages that are clear and simple, targeting both general public individuals and clinicians, and using television and radio [6,7]. Moreover, motivating

physicians to be involved in communicating with patients about the appropriate use of antibiotics and antibiotic resistance is also important [7]. Physicians' participation in developing campaign messages and communicating with the general public might significantly improve the chances of success. Early engagement can also have an impact on the sense of ownership of the campaign and facilitate physicians' consistent support [27]. This may indirectly influence physicians, focusing their attention on antibiotic prescribing and providing greater patient education.

Variable Sociological Factors

Awareness messages that display local surveillance data or amount of antibiotics locally misused may be important. Sharing with the general public the real-life experiences of individuals who have been infected with "superbugs" could be useful and may help the audience identify with those affected. Sharing real-life medical experience has been shown to be a useful communication platform from which to clarify public health-related stories, such as acne and cancer [28,29]. Considering these factors when designing awareness campaigns about antibiotics may help make an impact in GCC communities. Based on the Health Belief Model, perceived susceptibility (ie, you are at risk of getting infected) can be used to raise awareness [30,31]. However, delivering known messages to the target audience might result in a loss of interest and later disengagement. Educational interventions would be more successful if local contexts and barriers are adequately analyzed and addressed.

Replicating campaign strategies that have been initiated in different global regions outside the GCC region might not result in an effective outcome. It is crucially important to study the cultural factors and antibiotic distribution infrastructure in the GCC before thinking about the awareness messages. For example, the United States' Get Smart campaign [32] highly recommends that parents do not demand antibiotics for their children from the treating physician. We do not believe that this message line will be as effective as it might be in the United States because antibiotics can be purchased without a prescription from community pharmacies; therefore, a doctor's refusal might not make a difference. Approximately 37% of the total population of the GCC states consists of nonnational expatriates [33]; hence, it is important to consider cultural differences and not to neglect this segment when setting up a public awareness campaign. For example, the Saudi Ministry of Health has generated educational materials in multiple languages to fulfill this requirement [34].

Community pharmacies have a significant role in dispensing antibiotics in GCC communities. For example, 24.4% of 1645 recently surveyed antibiotic transactions in community pharmacies in Abu Dhabi were carried out without a prescription, including amoxicillin-clavulanic acid for sore throats and ceftriaxone for sexually transmitted infections [35]. The illegal practice of selling antibiotics over-the-counter, without a prescription, did not favor expatriates over citizens in the surveyed pharmacies in Abu Dhabi [15].

The self-prescription of antibiotics is another sociological factor that must be considered when designing antibiotic awareness campaigns in the GCC states [36,37]. This factor is strongly associated with the availability of antibiotics over-the-counter.

Before creating content to be used for awareness campaigns, it is necessary to conduct formative research to assess the public's existing knowledge of antibiotics resistance, understand the motivations for inappropriate antibiotics use, and learn about the social and cultural backgrounds for the targeted population. That will subsequently help to develop tailored key messages that can potentially encourage behavioral change [30]. Knowing these critical elements has lead to the success of many awareness campaigns, such as The Magic Glasses video to prevent soil-transmitted helminthes in China [31,38]. On the other hand, content produced for social media-based campaigns can be unrelated to the campaign's target. For example, it was found that the majority of Movember campaign-related tweets did not associate with prostatic and testicular cancer awareness [39], and the majority of tweets produced during breast cancer awareness week did not promote any specific preventive behaviors [40]. Despite the importance of developing related content, research into the correlation between social media-based awareness campaigns and behavior change is minimal because it is a new avenue in public health awareness [41].

Social media platforms can also contain contradictory health messages with potentially negative impact. Because social media platforms give users the freedom to publish their content, some of that content can contain medically misleading information, as found in YouTube videos promoting anorexia [42].

Funding

Funding is an important factor that may significantly affect a campaign's functionality and outcome. A systematic review of more than 20 international campaigns aimed at raising awareness of antibiotic use showed that these campaigns sourced their funding from different sectors, including the pharmaceutical industry [6]. The funding spent by pharmaceutical companies on promoting and marketing antibiotics is massive. For example, in 1998, it was estimated that pharmaceutical companies in the United States spent approximately US $1.6 billion to promote antibiotics [43]. On the other hand, media campaigns that encourage the prudent use of antibiotics are not widely supported [6]. Government funding is important. Because antibiotic awareness campaigns might translate into wiser use of antibiotics and potentially lead to a reduced selection of resistant bacteria, public funding should be offered to support awareness campaigns.

The cost of running a traditional mass media campaign to promote prudent antibiotic use in the community can be very expensive. For example, developing and conducting the French antibiotic awareness campaign carried out from 2002 to 2007 cost approximately €500 million over a 6-year period [44], whereas the Belgian campaign cost considerably less at approximately €400,000 per year [21]. The "Get Smart Colorado" campaign, which took place for 4 months in 2002, reported a cost of US $88,500 to purchase advertising space that included bus tails, bus stop posters, interior bus signs, and national public radio spots [10]. Similarly, the recent Italian campaign in 2013 cost approximately US $60,800 for purchasing media spots on television and radio, and in newspapers.

Approximately the same cost was spent to develop and print written visual aid materials, such as posters [9].

Considering social media platforms are free, establishing a social media-based campaign may be far cheaper than traditional media-based campaigns. However, in order to maintain continuous cyber presence and followers scalability, social media managers are usually hired [45], which can be an additional cost burden on social media-based campaigns. The key advantage of social media is the possibility to measure and track impressions and responses to online posts. These data can be used to guide social media campaigns to improve marketing strategy. However, platforms available to analyze big data generated from social media can be costly and may require technical expertise.

Time

Repeating the educational intervention over a long period of time is essential for the awareness success of mass media campaigns. Repetition over a long period of time has been demonstrated for causes such as smoking cessation and has helped achieve effectiveness [46]. The vast majority of antibiotic awareness campaigns launched in high-income countries between 1990 and 2007 were conducted over a period of more than 1 year [6]. Other campaigns, such as European Antibiotic Awareness Day (on November 18 each year) [47], are seasonal and have a long-term sustainable plan. However, the "Get Smart Colorado" campaign, which lasted for only 4 months, successfully showed a 3.8% net reduction in antibiotic dispensing at retail pharmacies as well as an 8.8% net decrease in managed care-associated antibiotic dispensing [10].

The ease of using social media, along with the indirect community contribution via "share" and "retweet" features, might provide long-term exposure and awareness messages to the wider general public. However, it is important to consider the temporal effect of social media feeds due to their short lifespan. It was found that the half-life of a tweet is approximately 24 minutes, whereas the half-life of Facebook posts is approximately 90 minutes [48]. A hashtag is a keyword preceded by a hash sign (#) that is used to identify and categorize messages on a specific topic, which can give the topic a longer lifespan in social media [49,50]. Keeping the audience engaged and interested in the topic is another important consideration. This might be achieved by ensuring that the key messages and materials are not overrepeated throughout the campaign's duration. Updating campaign materials with new and relevant data might keep the audience more engaged and keen to receive updated educational materials. Lastly, the time chosen to post the social media message can be critical for the lifespan of social media posts [51].

The Value of Social Media Platforms to Communities in the Gulf Cooperation Council States

The total number of users of the social networking website Facebook in the Arab world (22 countries) had grown to 54,552,875 by the end of May 2013; 33.4% of users are female and 68% are younger than 30 years [52]. Facebook users in the GCC states represent approximately 22% (12 million) of the total Facebook users in the Arab world [52] (Table 2).

Table 2. The use of Internet and social media platforms in the GCC states.

Country	Population (million)[a]	Internet users (million), n (%)[b]	Facebook users (million),[c] n (%)	Twitter users (million),[d] n (%)
Saudi Arabia	28.4	13.0 (45.8)	6.4 (22.5)	1.9 (6.7)
United Arab Emirates	8.3	5.7 (71.0)	3.4 (41.7)	0.4 (4.8)
Kuwait	3.1	2.0 (63.2)	0.8 (26.8)	0.2 (7.3)
Bahrain	1.2	1.0 (80.0)	0.3 (25.1)	0.1 (5.6)
Qatar	1.7	1.7 (99.9)	0.6 (34.4)	0.1 (4.4)
Oman	3.3	2.1 (63.6)	0.5 (16.4)	0.04 (1.2)

[a] Population figures obtained from [53].

[b] Internet user figures obtained from [54].

[c] Facebook user figures obtained from [52].

[d] Twitter user figures obtained from [55].

For the microblogging website Twitter, the number of active users in the Arab world reached 3,766,160 individuals as of March 2013, with an estimated 10,832,000 tweets per day. Saudi Arabia has the highest number of active Twitter users in the Arab world, with 1.9 million individuals, which is approximately 50% of the total Twitter users in the Arab region. Approximately 47% and 11% of the total tweets in the Arab world are generated from Saudi Arabia and United Arab Emirates, respectively.

The video-sharing website YouTube is also a popular media platform in the GCC, particularly in Saudi Arabia. As an update to research conducted by Forbes Middle East, we present data from selected local GCC talk shows on YouTube (Table 3). It is clear that these shows attract many viewers, although some shows from Saudi Arabia receive the most attention. This audience would make these shows an excellent platform for delivering awareness messages to a larger number of viewers.

Table 3. The popularity of selected YouTube-based shows in the GCC states.

Name of show	Origin	Launch date	Episodes[a]	Subscribers[a]	Total views[a]	Average views per episode[a]
EyshElly	Saudi Arabia	Feb 2011	63	1,714,699	197,686,128	3,137,875
3al6ayer	Saudi Arabia	Sep 2010	45	810,119	61,462,338	1,365,029
	Saudi Arabia	Sep 2010	33	649,465	72,887,865	2,208,723
Endam Cinema	Oman	Jul 2013	6	215	50,864	8477
Balalee6	United Arab Emirates	Jun 2012	13	6686	972,165	74,781
	United Arab Emirates	May 2012	5	10,133	954,815	190,963
shenoya3nitv	Kuwait	Jan 2012	54	67,477	8,076,193	149,559
How to Prevent from Corona	Saudi Arabia	May 2014	1	17,759	2,973,376	NA

[a] The figures were obtained from YouTube channels on November 15, 2013.

Examples of Saudi Public Health Awareness Messages Delivered Through Social Media Platforms

Owing to its high profile and popularity among Internet users in the GCC region, YouTube has often been used in Saudi Arabia as a platform to deliver public health–related awareness messages and campaigns (Table 4). We have noticed 2 different models for delivering health-related topics on YouTube in Saudi Arabia. One model was noticed in many campaigns that used comedy talk shows with large audience as a platform to deliver the awareness messages. With the help of other social media platforms, such as Facebook and Twitter, these messages have traveled far and wide, attracting a large number of viewers. For example, a Saudi-based comedy show named “Fe2aFala” released an episode about acquired immune deficiency syndrome (AIDS) and this episode attracted over 1 million viewers.

Table 4. Health messages delivered through Saudi YouTube-based shows.[a]

Show/channel name	Type	Awareness about	Number of viewers	Channel subscribers
Phosphine	Special episode	Phosphine gas	3,288,241[b]	14,821
Lumink	Special show	Health promotion	2,583,534	807,561
Telfaz11	Special episode	Breast cancer	1,489,802	296,563
Fe2aFala	Special episode	AIDS	1,139,948	351,907
Sen_tube	Entire show	Dental care	1,292,959	29,373
3almezan	Entire show	Obesity and well-being	1,097,713	51,746
Hotcoldshow	Special episode	Diabetes	455,485	81,725
MedScoope	Entire show	Health promotion	88,703	937

[a] The figures were obtained by accessing the YouTube channels on the March 4, 2014.

[b] This number of views was achieved within only 3 days of uploading the video on YouTube.

Another model is to use YouTube as a channel similar to traditional mass media to distribute health awareness messages. For example, in June 2014, the Saudi Arabia Ministry of Health launched an engaging public awareness campaign using YouTube, Twitter, Facebook, educational posters, and health guideline updates to educate the general public and medical communities on the emergence of, and health precautions needed for, MERS-CoV [34]. In this example, social media platforms may have been used to create online presence to endorse awareness messages delivered locally on traditional media. It must be noted that the incidence rate of MERS-CoV has declined [56] with the multi-faceted intervention, which social media has been a part of.

Publishing awareness content on social media might allow international distribution and an indefinite exposure period. However, in order to potentiate the effectiveness of a YouTube video, additional advertising strategies must take place. Table 4 summarizes some of the campaigns that have used YouTube shows to deliver public health awareness in Saudi Arabia and have attracted a high number of views.

Our Pilot Experience

Many health care facilities and organizations have started using Twitter as a teaching tool; for example, “tweeting” about specific health problems has been used by major organizations such as the Centers for Disease Control and Prevention (CDC), WHO, the National Cancer Institute (NCI), and the National Institutes of Health (NIH). “Tweets” go out from a sender and

are simultaneously received by all members of a group of "followers," providing a fast, open, and easy way to deliver a particular focused message.

We used Twitter in our online Arabic-language pilot campaign focusing on superbugs using a hashtag. We delivered short tweets and links to various articles and videos related to superbugs. We also translated multiple non-Arabic research articles and news for our more than 34,000 followers, who were primarily from Saudi Arabia and the surrounding areas. This medium gave us the ability to have a real-time conversation by answering their questions and concerns about the topic. To track our efforts, we regularly reviewed the number of followers, updates, retweets, and "mentions" in Twitter. Many evaluation metrics for Twitter can be collected. We evaluated our Twitter pilot study by analyzing the influence of some of our tweets. We kept track of how many of our "followers" published updates including "retweets" or "at replies" over time. A simple analysis of 147 selected tweets with the Arabic superbug hashtag resulted in approximately 4100 retweets between July and November 2013. This information does not tell us which of the tweets encouraged followers to go on to read the full article or watch the video, but it does give us useful information on general interest levels.

We created a whiteboard animation with a voiceover video [57] that discussed the importance of antibiotics, how superbugs are created, and the spread of superbugs in the GCC. We also discussed possible factors that could contribute to the emergence and spread of superbugs in addition to advice on how to control superbugs as reviewed previously [1].

Interestingly, the number of retweets that we received from our superbugs hashtag correlated with the number of YouTube views that we recorded for our video. The video viewing pattern as determined by YouTube was as expected: there were a large number of views soon after the video was posted and released on Twitter, and there were fewer views from 3 weeks later until the present day. We observed 2 large peaks in viewing rates within the first 2 weeks of the video being released. We believe the first one was due to our successful Twitter activities. We are unclear about what caused the second and largest jump in YouTube views. It is possible that the video was picked up by another group and spread on Twitter again (Figure 1).

We also observed sporadic smaller peaks in YouTube views from 3 weeks onward. These numbers suggest that the video was still drawing attention and being shared, most likely on social media, even months after being posted.

Our viewer retention rate was consistent with what YouTube estimates to be the average retention rate for videos of a similar length—approximately half of the video. This rate suggests that the length of our video—5.5 minutes—is adequate to capture the viewers' attention.

Subsequently, we created an infographic video uploaded on YouTube [58] that combined a selection of our tweets with visual and audio enhancement. The tweet that featured the video received 152 retweets and 130 favorites, and the video was viewed more than 3000 times in only 2 weeks (Figure 2).

The content generated for our pilot campaign was primarily dependent on our existing knowledge of antibiotic misuse in Saudi Arabia and neighboring states as previously reviewed [1,14-16,35-37]. We acknowledge the limitations of our pilot campaign resulting from the lack of formative research, its Arabic content, and reliance on social media. In order to achieve prospective behavioral change, it is important to understand the motivation for antibiotic misuse in the general public. Future research should conduct a thorough survey, interviews, and observations to cover the diverse population of the GCC states. This will help us design relevant key messages to be used in future awareness campaigns. The diverse ethnic groups and socioeconomic status of people in the Gulf States should also be considered in future campaigns because more languages and media platforms might need to be used. Relying solely on a social media platform in an awareness campaign might overlook the large population of migrant workers in the GCC who might not have access to social media platforms.

Figure 1. YouTube views over a period of 84 days for the whiteboard animated and voiceover video about superbugs and proper antibiotics use.

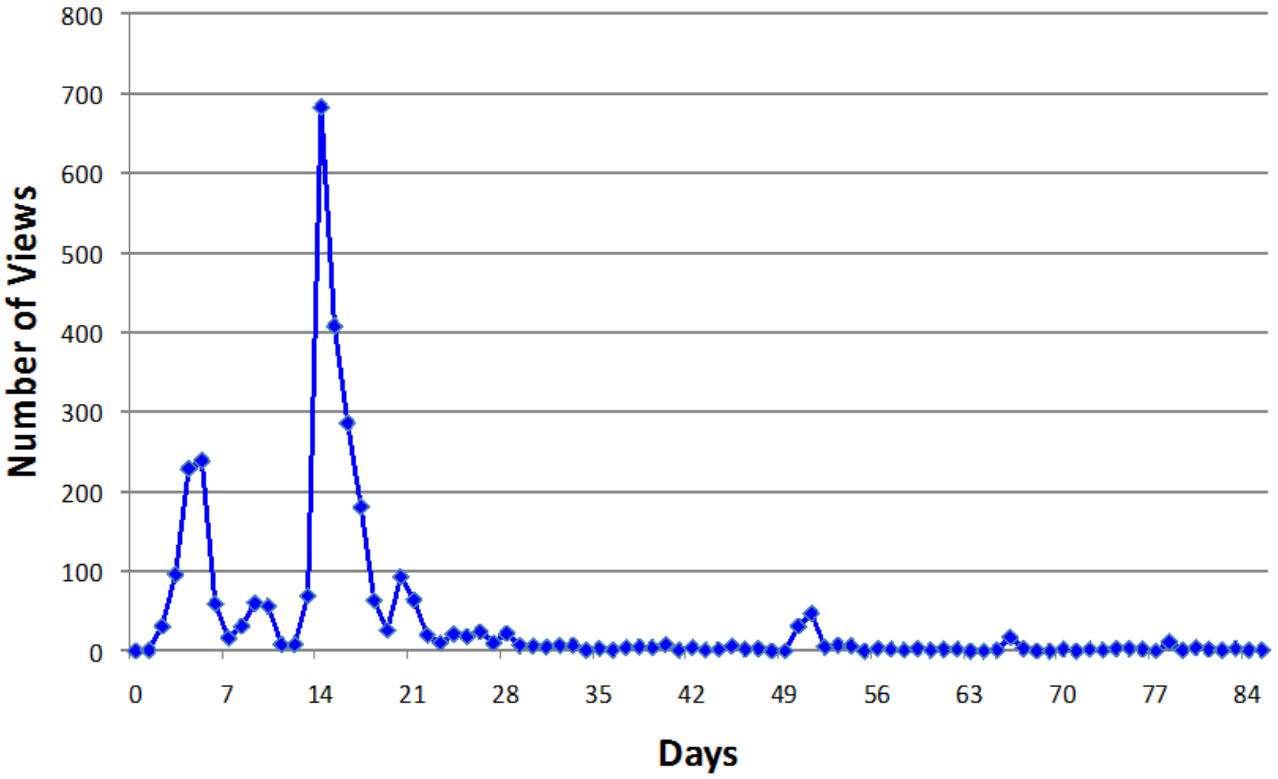
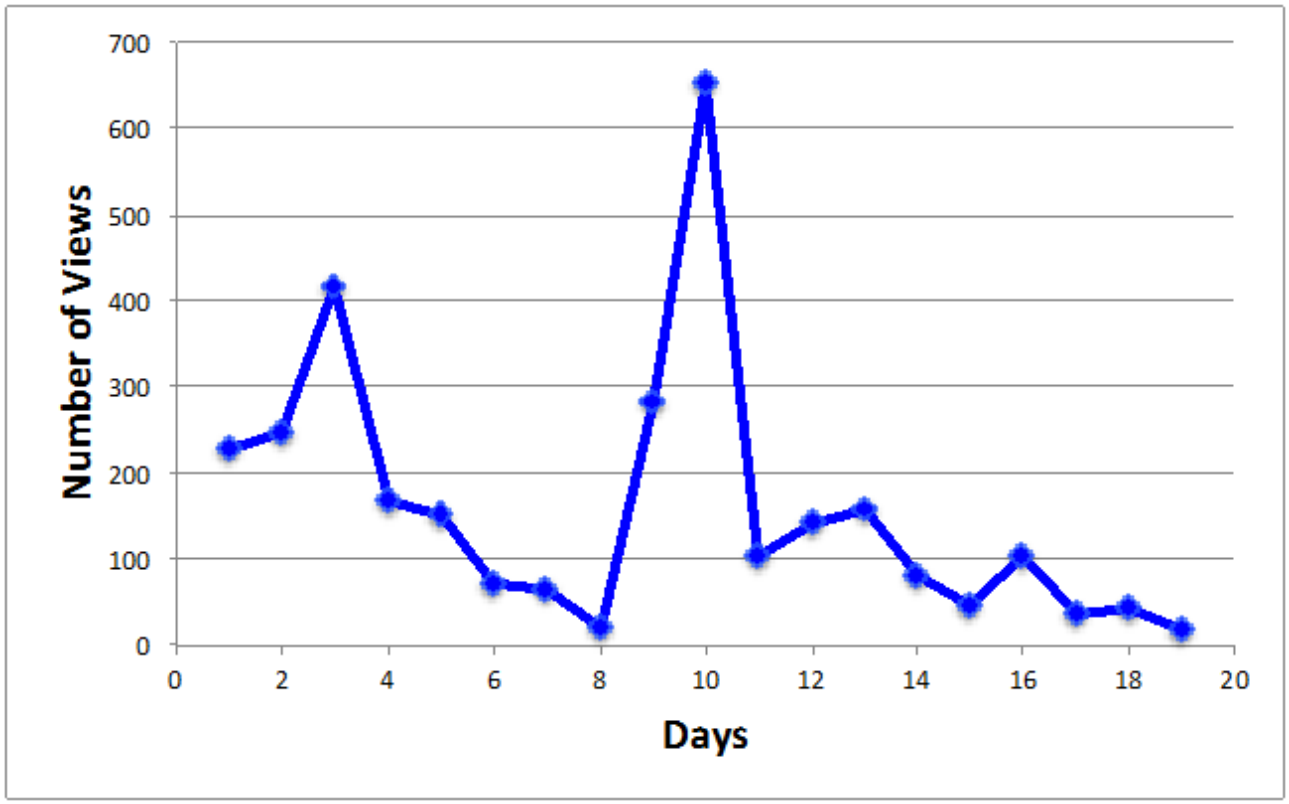


Figure 2. YouTube views over a period of 20 days for the infographic video.



Conclusion

Reducing the suboptimal use of antibiotics among the general public and medical community through awareness activities is an important element in national plans to combat rising AMR. However, it is important to create awareness content that is related to the target audience and based on formative research. Social media platforms seem to be a valuable platform for delivering awareness messages. Owing to social media popularity, awareness messages could reach a large number of

users and the reach can be tracked. Through our pilot experience we have successfully distributed antibiotics awareness messages through Twitter and YouTube to our target audience in the Gulf counties and Saudi Arabia. The use of social media can also enhance awareness campaigns delivered in traditional media channels. However, it is important to consider the cultural demographic diversity, which could limit the reach of awareness massages, such as the high population of immigrant workers in the GCC who might not have access to these emerging social media platforms. Social media–based messages can also have short life span, which might limit the effect and reach of key awareness messages. Measuring the impact of the social media–based awareness campaigns on the general public's understanding and behavioral change is a challenge and needs further research.

Acknowledgments

HMZ is academically sponsored by the government of Saudi Arabia to pursue postgraduate studies in the field of clinical microbiology and infectious diseases.

Conflicts of Interest

None declared.

References

1. Zowawi H, Balkhy H, Walsh T, Paterson D. β-Lactamase production in key gram-negative pathogen isolates from the Arabian Peninsula. Clin Microbiol Rev 2013 Jul;26(3):361-380 [FREE Full text] [doi: 10.1128/CMR.00096-12] [Medline: 23824364]
2. World Health Organization. 2011. World Health Day - 7 April 2011 URL: http://www.who.int/world-health-day/2011/en/ [accessed 2014-09-05] [WebCite Cache ID 6SLmsfwUh]
3. Lipsitch M, Samore M. Antimicrobial use and antimicrobial resistance: a population perspective. Emerg Infect Dis 2002 Apr;8(4):347-354 [FREE Full text] [doi: 10.3201/eid0804.010312] [Medline: 11971765]
4. Mauldin P, Salgado C, Hansen I, Durup D, Bosso J. Attributable hospital cost and length of stay associated with health care-associated infections caused by antibiotic-resistant gram-negative bacteria. Antimicrob Agents Chemother 2010 Jan;54(1):109-115 [FREE Full text] [doi: 10.1128/AAC.01041-09] [Medline: 19841152]
5. Huttner B, Goossens H, Verheij T, Harbarth S. Characteristics and outcomes of public campaigns aimed at improving the use of antibiotics in outpatients in high-income countries. Lancet Infect Dis 2010 Jan;10(1):17-31. [doi: 10.1016/S1473-3099(09)70305-6] [Medline: 20129146]
6. Albrich W, Monnet D, Harbarth S. Antibiotic selection pressure and resistance in Streptococcus pneumoniae and Streptococcus pyogenes. Emerg Infect Dis 2004 Mar;10(3):514-517 [FREE Full text] [doi: 10.3201/eid1003.030252] [Medline: 15109426]
7. McNulty CA, Nichols T, Boyle P, Woodhead M, Davey P. The English antibiotic awareness campaigns: did they change the public's knowledge of and attitudes to antibiotic use? J Antimicrob Chemother 2010 Jul;65(7):1526-1533 [FREE Full text] [doi: 10.1093/jac/dkq126] [Medline: 20488985]
8. World Health Organization. 2011. World Health Day 2011: policy briefs URL: http://www.who.int/world-health-day/2011/presskit/whd2011_fs4_animal.pdf?ua=1 [accessed 2014-09-05] [WebCite Cache ID 6SLmbco8t]
9. Formoso G, Paltrinieri B, Marata A, Gagliotti C, Pan A, Moro M, et al. Feasibility and effectiveness of a low cost campaign on antibiotic prescribing in Italy: community level, controlled, non-randomised trial. BMJ 2013;347:f5391 [FREE Full text] [Medline: 24030722]
10. Gonzales R, Corbett K, Wong S, Glazner J, Deas A, Leeman-Castillo B, et al. "Get smart Colorado": impact of a mass media campaign to improve community antibiotic use. Med Care 2008 Jun;46(6):597-605. [doi: 10.1097/MLR.0b013e3181653d2e] [Medline: 18520314]
11. Balkhy H, El-Saed A, Jabri F. Antimicrobial consumption in four different adult intensive care units in a Saudi tertiary care hospital. 2011 Presented at: The Eleventh Congress of the International Federation of Infection Control (IFIC); Oct 2011; Venice, Italy.
12. National Nosocomial Infections Surveillance System. National Nosocomial Infections Surveillance (NNIS) System Report, data summary from January 1992 through June 2004, issued October 2004. Am J Infect Control 2004 Dec;32(8):470-485. [doi: 10.1016/S0196655304005425] [Medline: 15573054]
13. Vojtová V, Kolár M, Hricová K, Uvízl R, Neiser J, Blahut L, et al. Antibiotic utilization and Pseudomonas aeruginosa resistance in intensive care units. New Microbiol 2011 Jul;34(3):291-298 [FREE Full text] [Medline: 21811750]
14. Bin Abdulhak AA, Altannir M, Almansor M, Almohaya M, Onazi A, Marei M, et al. Non prescribed sale of antibiotics in Riyadh, Saudi Arabia: a cross sectional study. BMC Public Health 2011;11:538 [FREE Full text] [doi: 10.1186/1471-2458-11-538] [Medline: 21736711]
15. Dameh M, Green J, Norris P. Over-the-counter sales of antibiotics from community pharmacies in Abu Dhabi. Pharm World Sci 2010 Oct;32(5):643-650. [doi: 10.1007/s11096-010-9418-5] [Medline: 20640511]

16. Alumran A, Hou X, Sun J, Yousef A, Hurst C. Assessing the construct validity and reliability of the Parental Perception on Antibiotics (PAPA) scales. BMC Public Health 2014;14:73 [FREE Full text] [doi: 10.1186/1471-2458-14-73] [Medline: 24456730]
17. GCC Centre for Infection Control. The GCC Strategic Plan for Combating Antimicrobial Resistance. Saudi Arabia: The Ministry of National Guard, Health Affairs; 2015.
18. European Antibiotic Awareness Day. 2015. Toolkit for engaging in social media activities promoting prudent antibiotic use URL: http://ecdc.europa.eu/en/eaad/pages/social-media-guidance.aspx [accessed 2015-06-05] [WebCite Cache ID 6Z3LD4loN]
19. George KS, Roberts CB, Beasley S, Fox M, Rashied-Henry K, Brooklyn Partnership to Drive Down Diabetes (BP3D). Our health is in our hands: a social marketing campaign to combat obesity and diabetes. Am J Health Promot 2015 May 14:1. [doi: 10.4278/ajhp.130625-ARB-323] [Medline: 25973964]
20. Lambert DN, Bishop LE, Guetig S, Frew PM. A formative evaluation of social media campaign to reduce adolescent dating violence. JMIR Res Protoc 2014;3(4):e64 [FREE Full text] [doi: 10.2196/resprot.3546] [Medline: 25405494]
21. Goossens H, Coenen S, Costers M, De Corte S, De Sutter A, Gordts B, et al. Achievements of the Belgian Antibiotic Policy Coordination Committee (BAPCOC). Euro Surveill 2008 Nov 13;13(46):19036 [FREE Full text] [Medline: 19021954]
22. Sabuncu E, David J, Bernède-Bauduin C, Pépin S, Leroy M, Boëlle PY, et al. Significant reduction of antibiotic use in the community after a nationwide campaign in France, 2002-2007. PLoS Med 2009 Jun 2;6(6):e1000084 [FREE Full text] [doi: 10.1371/journal.pmed.1000084] [Medline: 19492093]
23. NPS MedicineWise. 2014. Australia must get wise about antibiotics to reduce the spread of superbugs Antibiotic Awareness Week November URL: http://www.nps.org.au/media-centre/media-releases/repository/Australia-must-get-wise-about-antibiotics-to-reduce-the-spread-of-superbugs-Antibiotic-Awareness-Week-12-18-November [accessed 2014-09-05] [WebCite Cache ID 6SLntTu8s]
24. McNulty CA, Lecky DM, Farrell D, Kostkova P, Adriaenssens N, Koprivová HT, e-Bug Working Group. Overview of e-Bug: an antibiotic and hygiene educational resource for schools. J Antimicrob Chemother 2011 Jun;66 Suppl 5:v3-12 [FREE Full text] [doi: 10.1093/jac/dkr119] [Medline: 21680584]
25. e-Bug Project Information. 2015. URL: http://www.e-bug.eu/partners/index.html [accessed 2015-06-05] [WebCite Cache ID 6Z3MrFsFG]
26. e-Bug Project Information. 2015. Web analytics URL: http://www.e-bug.eu/partners/partner_sa/sa.html [accessed 2015-06-05] [WebCite Cache ID 6Z3MzBS5Z]
27. LOcal Campaign on Antibiotics ALliance (LOCAAL) study group. Doctors and local media: a synergy for public health information?: a controlled trial to evaluate the effects of a multifaceted campaign on antibiotic prescribing (protocol). BMC Public Health 2011;11:816 [FREE Full text] [doi: 10.1186/1471-2458-11-816] [Medline: 22011332]
28. Chou W, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. J Med Internet Res 2011;13(1):e7 [FREE Full text] [doi: 10.2196/jmir.1569] [Medline: 21247864]
29. Vance K, Howe W, Dellavalle R. 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]
30. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
31. Bieri FA, Yuan LP, Li YS, He YK, Bedford A, Li RS, et al. Development of an educational cartoon to prevent worm infections in Chinese schoolchildren. Infect Dis Poverty 2013;2(1):29 [FREE Full text] [doi: 10.1186/2049-9957-2-29] [Medline: 24289667]
32. Centers for Disease Control and Prevention. 2014. Get Smart programs and observances URL: http://www.cdc.gov/getsmart/ [accessed 2014-09-05] [WebCite Cache ID 6SLnJLSaw]
33. Kapiszewski A. Arab versus Asian migrant workers in the GCC countries. 2006 May 22. URL: http://www.un.org/esa/population/meetings/EGM_Ittmig_Arab/P02_Kapiszewski.pdf [accessed 2014-09-02] [WebCite Cache ID 6SHg4DY29]
34. Saudi Ministry of Health, Command & Control Center. 2014. Important information about MERS-CoV URL: http://www.moh.gov.sa/en/CCC/aboutcorona/pages/default.aspx [accessed 2014-09-05] [WebCite Cache ID 6SLmxftfm]
35. Abasaeed A, Vlcek J, Abuelkhair M, Andrajati R, Elnour A. A comparative study between prescribed and over-the-counter antibiotics. Saudi Med J 2013 Oct;34(10):1048-1054. [Medline: 24145940]
36. Abobotain A, Sheerah H, Alotaibi F, Joury A, Mishiddi R, Siddiqui A, et al. Socio-demographic determinants of antibiotic misuse in children. A survey from the central region of Saudi Arabia. Saudi Med J 2013 Aug;34(8):832-840. [Medline: 23974456]
37. Abasaeed A, Vlcek J, Abuelkhair M, Kubena A. Self-medication with antibiotics by the community of Abu Dhabi Emirate, United Arab Emirates. J Infect Dev Ctries 2009;3(7):491-497 [FREE Full text] [Medline: 19762966]
38. Bieri FA, Gray DJ, Williams GM, Raso G, Li YS, Yuan L, et al. Health-education package to prevent worm infections in Chinese schoolchildren. N Engl J Med 2013 Apr 25;368(17):1603-1612. [doi: 10.1056/NEJMoa1204885] [Medline: 23614586]

39. Bravo CA, Hoffman-Goetz L. Tweeting about prostate and testicular cancers: what are individuals saying in their discussions about the 2013 Movember Canada campaign? J Cancer Educ 2015 Apr 24:e1. [doi: 10.1007/s13187-015-0838-8] [Medline: 25903054]
40. Thackeray R, Burton SH, Giraud-Carrier C, Rollins S, Draper CR. Using Twitter for breast cancer prevention: an analysis of breast cancer awareness month. BMC Cancer 2013;13:508 [FREE Full text] [doi: 10.1186/1471-2407-13-508] [Medline: 24168075]
41. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
42. Syed-Abdul S, Fernandez-Luque L, Jian WS, Li YC, Crain S, Hsu MH, et al. Misleading health-related information promoted through video-based social media: anorexia on YouTube. J Med Internet Res 2013;15(2):e30 [FREE Full text] [doi: 10.2196/jmir.2237] [Medline: 23406655]
43. Ma J, Stafford R, Cockburn I, Finkelstein S. A statistical analysis of the magnitude and composition of drug promotion in the United States in 1998. Clin Ther 2003 May;25(5):1503-1517. [Medline: 12867225]
44. Huttner B, Harbarth S. "Antibiotics are not automatic anymore"--the French national campaign to cut antibiotic overuse. PLoS Med 2009 Jun 2;6(6):e1000080 [FREE Full text] [doi: 10.1371/journal.pmed.1000080] [Medline: 19492081]
45. Albarran A. Management of Electronic and Digital Media. Boston, MA: Wadsworth, Cengage Learning; 2013. URL: http://www.cengagebrain.com.au/content/9781133824640.pdf [accessed 2015-06-05] [WebCite Cache ID 6Z3Ogq4Jn]
46. Bala M, Strzeszynski L, Cahill K. Mass media interventions for smoking cessation in adults. Cochrane Database Syst Rev 2008(1):CD004704. [doi: 10.1002/14651858.CD004704.pub2] [Medline: 18254058]
47. Earnshaw S, Monnet D, Duncan B, O'Toole J, Ekdahl K, Goossens H, European Antibiotic Awareness Day Technical Advisory Committee, European Antibiotic Awareness Day Collaborative Group. European Antibiotic Awareness Day, 2008 - the first Europe-wide public information campaign on prudent antibiotic use: methods and survey of activities in participating countries. Euro Surveill 2009 Jul 30;14(30):19280 [FREE Full text] [Medline: 19643056]
48. Rey B. Wisemetrics. 2014 Mar 05. Your tweet half-life is 1 billion times shorter than Carbon-14's URL: http://blog.wisemetrics.com/tweet-isbillion-time-shorter-than-carbon14/ [accessed 2015-06-05] [WebCite Cache ID 6Z3OrpXjg]
49. Glasgow K, Fink C. Hashtag lifespan and social networks during the london riots. In: Social Computing, Behavioral-Cultural Modeling and Prediction. Heidelberg: Springer Berlin; 2013.
50. Tsur O, Rappoport A. What's in a hashtag?: content based prediction of the spread of ideas in microblogging communities. In: Proceedings of the fifth ACM international conference on Web search and data mining. 2012 Presented at: WSDM '12; Feb 8-12, 2012; Seattle, WA p. 643-652. [doi: 10.1145/2124295.2124320]
51. Cooper BB. buffersocial. 2013 Dec 11. How to find your best time to tweet: the 4 most accurate methods URL: https://blog.bufferapp.com/the-4-most-accurate-ways-to-find-your-best-time-to-tweet [accessed 2015-06-05] [WebCite Cache ID 6Z3PNYfku]
52. Arab Social Media Report. 2014. Facebook in the Arab Region URL: http://www.arabsocialmediareport.com/Facebook/LineChart.aspx?&PriMenuID=18&CatID=24&mnu=Cat [accessed 2014-09-05] [WebCite Cache ID 6SLxXm0z9]
53. The Cooperation Council for the Arab States of the Gulf. 2012. URL: http://www.gcc-sg.org/eng/indexc64c.html?action=GCC [accessed 2014-09-05] [WebCite Cache ID 6SLo3D6q3]
54. Internet World Stats. 2013. Middle East URL: http://www.internetworldstats.com/middle.htm [accessed 2014-09-05] [WebCite Cache ID 6SLxhlNJ0]
55. Arab Social Media Report. 2014. Twitter in the Arab Region URL: http://www.arabsocialmediareport.com/Twitter/LineChart.aspx?&PriMenuID=18&CatID=25&mnu=CatN%20million%20%28%%29 [accessed 2014-09-05] [WebCite Cache ID 6SLxTUYs0]
56. Chan JF, Lau SK, To KK, Cheng VC, Woo PC, Yuen KY. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. Clin Microbiol Rev 2015 Apr;28(2):465-522. [doi: 10.1128/CMR.00102-14] [Medline: 25810418]
57. YourekaScience. YouTube. 2013 Oct 25. Arabic: because you care about me-Antibiotic resistance in GCC countries URL: https://www.youtube.com/watch?v=nedLlLDE4rU [accessed 2014-09-05] [WebCite Cache ID 6SLnQvGaW]
58. Medscoope. YouTube. 2014 Jan 26. [Be smart and know when to use antibiotics] URL: https://www.youtube.com/watch?v=dHmkmd78Dq4 [accessed 2014-09-05] [WebCite Cache ID 6SLni8oML]

Abbreviations

AIDS: acquired immune deficiency syndrome
AMR: antimicrobial resistance
GCC: Gulf Cooperation Council
ICU: intensive care unit
NCI: National Cancer Institute
NIH: National Institutes of Health

WHO: World Health Organization

Edited by G Eysenbach; submitted 24.09.14; peer-reviewed by B Dunais, E Castro-Sánchez; comments to author 22.04.15; revised version received 05.06.15; accepted 21.06.15; published 15.10.15.

Please cite as:
Zowawi HM, Abedalthagafi M, Mar FA, Almalki T, Kutbi AH, Harris-Brown T, Harbarth S, Balkhy HH, Paterson DL, Hasanain RA The Potential Role of Social Media Platforms in Community Awareness of Antibiotic Use in the Gulf Cooperation Council States: Luxury or Necessity?
J Med Internet Res 2015;17(10):e233
URL: http://www.jmir.org/2015/10/e233/
doi:10.2196/jmir.3891
PMID:26471079

Original Paper

Using WhatsApp and Facebook Online Social Groups for Smoking Relapse Prevention for Recent Quitters: A Pilot Pragmatic Cluster Randomized Controlled Trial

Yee Tak Derek Cheung[1], PhD; Ching Han Helen Chan[2], MPhil; Chi-Keung Jonah Lai[2], MSW; Wai Fung Vivian Chan[3], MPhil; Man Ping Wang[3], PhD; Ho Cheung William Li[3], PhD; Sophia Siu Chee Chan[3], PhD; Tai-Hing Lam[1], MBBS

[1]School of Public Health, The University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Integrated Centre on Smoking Cessation, Tung Wah Group of Hospitals, Hong Kong, China (Hong Kong)

[3]School of Nursing, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Yee Tak Derek Cheung, PhD
School of Public Health
The University of Hong Kong
5/F William MW Mong Block
21 Sassoon Road, Hong Kong
Hong Kong
China (Hong Kong)
Phone: 852 39176652
Fax: 852 2855 9528
Email: derekcheung@hku.hk

Abstract

Background: Quit attempters often have episodes of smoking relapse before they eventually quit. Interactive text messaging through mobile phones has been shown to increase abstinence. This service can be potentially applied on the platform of a social networking service to help quitters maintain abstinence.

Objective: Our aim was to determine if the group discussion and reminders via the WhatsApp or Facebook social group were effective to prevent smoking relapse in quitters who had stopped smoking recently.

Methods: This was a single-blinded, parallel, 3-arm pilot cluster randomized controlled trial allocating recent quitters, who had completed an 8-week treatment and reported abstinence for at least 7 days, to WhatsApp (n=42), Facebook (n=40), and a control group (n=54). The 2 intervention groups participated in a 2-month online group discussion with either WhatsApp or Facebook moderated by a trained smoking cessation counselor and received a self-help booklet on smoking cessation. The control group only received the booklet. The primary outcome was the 2- and 6-month relapse rates, defined as the proportion of participants who smoked at least 5 cigarettes in 3 consecutive days.

Results: Fewer participants in the WhatsApp group (17%, 7/42) reported relapse than the control group (42.6%, 23/54) at 2-month (OR 0.27, 95% CI 0.10-0.71) and 6-month (40.5%, 17/42 vs 61.1%, 33/54; OR 0.43, 95% CI 0.19-0.99) follow-ups. The Facebook group (30.0%, 12/40) had an insignificantly lower relapse rate than the control group (42.6%, 23/54) at 2-month (OR 0.58, 95% CI 0.24-1.37) and 6-month (52.5%, 13/40 vs 61.1%, 33/54; OR 0.70, 95% CI 0.31-1.61) follow-ups. The WhatsApp social groups had more moderators' posts (median 60, IQR 25 vs median 32, IQR 7; *P*=.05) and participants' posts (median 35, IQR 50 vs median 6, IQR 9; *P*=.07) than their Facebook counterparts, but the difference was insignificant.

Conclusions: The intervention via the WhatsApp social group was effective in reducing relapse probably because of enhanced discussion and social support. Inactive discussion in the Facebook social group might have attributed to the lower effectiveness.

ClinicalTrial: Clinicaltrials.gov NCT02007369; https://clinicaltrials.gov/show/NCT02007369 (Archived by WebCite® at http://www.webcitation.org/6c3RbltQG)

(J Med Internet Res 2015;17(10):e238) doi:10.2196/jmir.4829

KEYWORDS

social networking; social media; smoking cessation; relapse prevention

Introduction

The World Health Organization's MPOWER measures includes "offer help to quit tobacco use" as one of 6 effective tobacco control strategies [1]. Despite the availability of medication and counseling services, quit attempters often have smoking "slips" (ie, one or a few puffs) or relapses before sustaining longer abstinence [2]. Quitters who quit smoking recently have to manage nicotine withdrawal symptoms and smoking cues in their daily environment. Approximately one-third of quitters relapse smoking 3 months after completing smoking cessation treatment and this proportion is 50% for those who quit for a week or less [3]. A US longitudinal study of smokers who received smoking cessation pharmacotherapies in primary care clinics found that approximately 80% relapsed smoking within a year after the treatment [4].

A meta-analysis of relapse prevention interventions showed that smoking cessation drugs to reduce nicotine cravings and withdrawal symptoms, such as bupropion (pooled OR 1.49, 95% CI 1.10-2.01) and nicotine replacement therapy (NRT) (pooled OR 1.33, 95% CI 1.08-1.63) were effective in preventing smoking relapse for at least 12 months [5]. However, the prevalence of use was low [6] and the compliance was poor (ie, use for less than 8 weeks) [7]. Previous randomized controlled trials (RCTs) using individual counseling or self-help written materials had inadequate sample size to support the effectiveness of relapse prevention [8]. Group counseling was effective for preventing relapse in the short term (eg, 3 months) (pooled OR 2.55, 95% CI 1.58-4.11), but the effect dissipated at long-term follow-up [5]. The limited effect might be explained by only a few face-to-face group sessions [9-11], which failed to offer instant and continuing support for recent quitters to manage craving or smoking cues.

Mobile phone-based interventions are potentially effective to support recent quitters to quit [12,13] and prevent relapse [14,15]. In Hong Kong, with approximately 7 million residents, there are more than 17 million subscribers to mobile phone services and more than 12 million of them are 2.5G/3G/4G subscribers with mobile Internet services [16]. Because Internet access with mobile phones has become popular, interventions via social networking services to support health-related behavior change have been examined for weight control and increasing physical activity [17]. A systematic review showed that such interventions have small to moderate effect size (-0.05 to 0.84), with only 2 of 7 studies showing statistically significant effects [17]. A few exploratory studies showed that a social networking service enabled reaching a sizable number of smokers in the community and increased peer interaction [18-21]. It can be a platform for smokers who seek immediate assistance and professional advice when they need it [22].

In this pilot RCT, we tested the effectiveness of a relapse prevention intervention using WhatsApp and Facebook, 2 common mobile phone apps in Hong Kong, to reduce smoking relapse in recent quitters who had just completed treatment at smoking cessation clinics.

Methods

Trial Design

The pilot single-blinded, pragmatic, parallel 3-armed cluster RCT compared the relapse rate at 2- and 6-month follow-ups between recent quitters who participated in group discussion and received reminders (group A: WhatsApp; group B: Facebook) and those who did not (group C: control; allocation ratio 1:1:1). The study was approved by the Institutional Review Board of the University of Hong Kong / Hong Kong Authority Hong Kong West Cluster (IRB reference no: UW-13-528).

Participants

All participants were clients of the Tung Wah Group of Hospitals Integrated Centre of Smoking Cessation (ICSC) in China Hong Kong, which provides 8-week free treatment, including counseling, telephone follow-ups, physicians' assessment, and prescription of free NRT or varenicline (a smoking cessation drug to relieve cravings while blocking the reinforcing effects of nicotine) [23]. At 8-week follow-up during telephone or face-to-face counseling, clients were asked by the ICSC counselors if they had quit. Self-reported quitters were then screened with the criteria for eligibility, including (1) daily smoker at first entry to the ICSC, (2) aged 18 years or older, (3) received 3 to 8 smoking cessation counseling sessions provided by the ICSC, (4) reported tobacco abstinence for at least 7 days, (5) able to communicate in Cantonese and read and write Chinese, (6) had a mobile phone through a local network, and (7) were able to access the Internet by mobile phone. Clients were excluded who had unstable psychological conditions as advised by physicians, possible alcohol dependence as measured by the Alcohol Use Disorders Identification Test (AUDIT) [24], or were pregnant.

Interventions

Treatment conditions for groups A and B included participation in the WhatsApp or Facebook online social group, respectively, and a 22-page booklet related to quitting and healthy diet. The social group function of WhatsApp and Facebook was used as the intervention platform because it supports a real-time sharing of text and multimedia messages among group members. Each social group started on the first day after each recruitment week and closed after 2 months. Group members received 3 reminders per week from a moderator who was a social worker or nurse with experience in smoking cessation counseling. These reminders, including texts, pictures, and videos, were based on the "Treatments for the Recent Quitter" of the US Clinical Practice Guidelines on Treating Tobacco Use and Dependence [2], including (1) encourage to maintain abstinence, (2) remind about the importance of remaining abstinence, (3) prevent smoking triggers, (4) remind about the withdrawal symptoms and lapse, (5) advise about stress and mood management, and (6) advise about weight control (Multimedia Appendix 1). All

moderators were provided a guideline in sending reminders, enhancing discussion, and other tasks in the social group (Multimedia Appendix 2).

Participants' privacy was protected by advising them to change the privacy setting in their WhatsApp and Facebook accounts and setting up participation rules. Because telephone numbers appear in the WhatsApp social groups, male and female participants in the WhatsApp group were separated into different social groups to avoid the possibility of misconduct or harassment, which was a concern raised by some female respondents in our pilot qualitative interviews. Telephone numbers can be concealed in Facebook; therefore, sex separation was not applied. Group C was a control group; they received only the same self-help booklet and were advised to contact ICSC's counselors when they faced high-risk situations or had smoking lapses (usual care).

Follow-Up

All participants were contacted via telephone at 2- and 6-month follow-ups after the random group allocation. Interviewers were blinded to the group assignment. Participants who reported tobacco abstinence in the past 7 days were visited by trained staff to collect their exhaled carbon monoxide (CO) and saliva sample. The participant was given HK $100 (approximately US $12.80) if his/her exhaled CO was less than 4 parts per million (ppm) and salivary cotinine was less than 10 ng/mL, which confirmed abstinence [25,26]. To minimize the incentive effect, if any, on the validation result, the incentive was small, only enough to compensate for travel and a little time cost. All participants were unaware of the incentive before follow-up and only the participants who reported abstinence were notified of the incentive.

Outcomes

The primary outcome was self-reported relapse rate, which was defined as the proportion of participants who self-reported smoking at least 5 cigarettes in 3 consecutive days in the past 2 months at the 2-month follow-up [27]. Another primary outcome was the 4-month relapse rate at 6-month follow-up. Secondary outcomes included (1) self-reported any smoking incidence (ie, smoking lapse), (2) self-reported smoking in the past 7 days, and (3) biochemically validated abstinence at the 2 follow-ups.

The questionnaire also collected other smoking-related information, including frequency of smoking urges in the past week [28], intensity of smoking urge in the past 24 hours [29], thinking of enjoying smoking [28], the Minnesota Nicotine Withdrawal Scale (MNWS) [29], and the 12-item Smoking Self-Efficacy Questionnaire (SEQ-12) [30].

Sample Size

The expected sample size was 40 for each group (ie, total sample size=120) to generate preliminary estimates of the intervention effectiveness. ICSC's treatment record showed for quitters at the RCT enrollment, approximately one-third reported a smoking relapse at 6-month follow-up. Assuming the quit rate of group C was 33.3% and the effect size of the primary outcome between groups A/B and C was 1.5, the estimated relapse rates for groups A/B and C were approximately 22% and 33%, respectively. The power for detecting this difference in 120 participants using the Fisher exact test is 22%, suggesting we might wrongly accept the null hypothesis (ie, no difference between groups A/B and C).

Randomization

Cluster randomization was used to allocate all participants recruited in a particular week to one RCT group. This randomization could allocate a sufficient number of participants in a social group each week and the selection bias due to recruitment week was unlikely. The estimated recruitment period was 9 weeks and each week was randomized to group A, B, and C using numbers generated on a website for generating random variables (RANDOM.ORG) by one of the authors (CYTD). After the 9-week recruitment, the number of participants in groups B and C were only 19 and 27, respectively. Therefore, we extended the recruitment period by 5 additional weeks and used the same randomization method.

Concealment Mechanism

The ICSC counselors who screened and enrolled participants were notified of the group allocation on Monday of each recruitment week. Participants were not aware of the allocation sequence.

Blinding

All participants received a specific relapse prevention intervention, but they did not know what the other interventions were. All assessors of outcomes were blinded to the RCT group of each participant.

Statistical Analysis

Data were entered into SPSS for Windows version 20 for analysis. Descriptive statistics including frequency, percentage, and mean were used to summarize the outcomes and other variables. Chi-square tests were used to compare categorical variables between subgroups. The Kolmogorov-Smirnov test was used to determine the use of *t* test (normal distribution) or Mann-Whitney *U* test (nonnormal distribution) for the comparison. We analyzed the primary and secondary outcomes with Fisher exact test and odds ratios with and without adjustment for significantly different characteristics at baseline. By intention-to-treat (ITT) analysis [31], participants who were lost or refused to follow up were treated as having smoking relapse and lapse. Sensitivity analysis assuming that missing participants had not changed smoking status since the previous follow-up (last observation carried forward [LOCF]) and excluding participants who were lost to follow-up (complete-case analysis) were performed for the primary outcomes. Additional analysis excluding those in groups A and B who did not participate in the social groups was also conducted. General linear model repeated measures analysis was used to examine the changes of other smoking-related variables.

A content analysis of all the posts in the social groups was conducted to understand how the intervention helped participants prevent relapse. All posts in the WhatsApp and Facebook social groups were archived before the social groups were closed by

the moderator. Each post was coded by 2 researchers separately and was classified by their content. The Mann-Whitney *U* test was used to compare the median number of posts between the WhatsApp and Facebook social groups because we had no assumption about their statistical distribution.

Results

Group Allocation and Retention Rates at Follow-Ups

From February 2014 to May 2014, 247 quitters were screened for eligibility. Of these, 68 quitters (27.5%) were ineligible, 41 (16.6%) refused to participate, and 2 (0.8%) had incomplete intake information. In all, 136 quitters (55.1%) participated with 42 allocated to group A (WhatsApp), 40 to group B (Facebook), and 54 to group C (Control) (Figure 1). The major reasons for ineligibility were possible alcohol dependence measured by AUDIT (n=30), had no mobile phone (n=23), or no Internet access by mobile phone (n=23).

For the 136 participants, 86.8% (118/136) were successfully followed at 2-month follow-up with 88% (35/42) in group A, 95% (36/40) in group B, and 80% (43/54) in group C (Figure 1). The overall retention rate at 6-month follow-up was 73.5% (100/136), with 81% (34/42) in group A, 70% (28/40) in group B, and 70% (38/54) in group C. The reasons for loss to follow-ups were (1) unable to reach via telephone, (2) refusal to follow up, and (3) incomplete survey.

Demographic Characteristics and Smoking Profile

For the 136 participants, 76.5% (104/136) were male and the mean age was 40.5 (SD 9.9) years. The mean age of group B was significantly lower than group C (*P*=.01) (Table 1). There was a significant difference in the negative affect subscale of the MNWS between groups B and C (*P*=.02) and for insomnia between groups A and C (*P*=.01) (Table 2). Differences in other sociodemographic characteristics and smoking profile in the 3 RCT groups were not significant. More than 90% (128/136, 94.1%) had been prescribed free NRT in their smoking cessation treatment, but only 5 of 136 (3.7%) were prescribed varenicline (Table 2). There was also no significant difference in other smoking-related variables. Overall, 86.0% (117/136) of the participants had been abstinent for at least 28 days before joining the RCT with 81% (34/42) in group A, 95% (38/40) in group B, and 83% (45/54) in group C.

Lapse and Relapse Rates

In the ITT analysis, fewer participants in group A reported smoking relapse than in group C at 2-month (17%, 7/42 vs 43%, 23/54; OR 0.27, 95% CI 0.10-0.71 *P*=0.008; power=74.6%) and 6-month (41%, 17/42 vs 61%, 33/54; OR 0.43, 95% CI 0.19-0.99, *P*=.049; power=54.9%) follow-ups. Also, the odds ratios of 2-month relapse rate adjusting for baseline difference in smoking urge and days of abstinence were significant (adjusted OR 0.26, 95% CI 0.09-0.74, *P*=.01) (Table 3). There was no significant difference in the relapse rate between groups B and C at 2-month (30%, 12/40 vs 43%, 23/54; OR 0.58, 95% CI 0.24-1.37, *P*=.21; power=36.9%) and 6-month (53%, 21/40 vs 61%, 33/54; OR 0.70, 95% CI 0.31-1.61, *P*=.40; power=20.6%) follow-ups. The power analysis showed that the comparison of the relapse rate between groups B and C had a large type II error (ie, accepting the null hypothesis when it was false). The odds ratios comparing 2-month self-reported relapse rate between groups A and C using ITT, LOCF (assumed to have the same smoking status as the last follow-up), and complete-case analysis were mostly significant and consistent.

Figure 1. CONSORT flow diagram of the pilot randomized controlled trial. AUDIT: The Alcohol Use Disorders Identification Test.

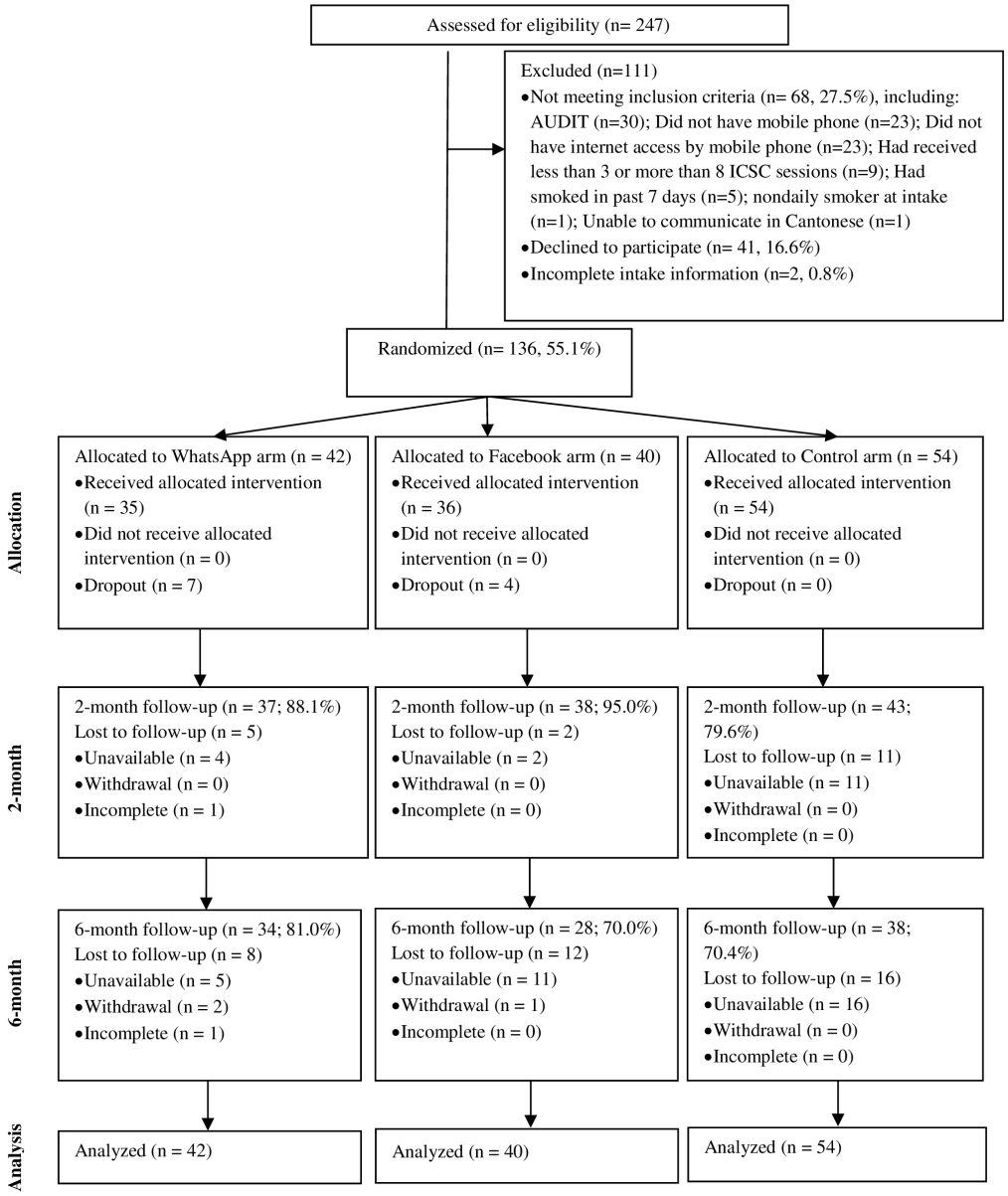


Table 1. Sociodemographic and smoking characteristics of participants at entry to the smoking cessation clinics (N=136).

Baseline characteristics	Group A: WhatsApp n=42	Group B: Facebook n=40	Group C: control n=54	*P* [a]	
				A vs C	B vs C
Gender, n (%)					
Male	32 (76)	29 (73)	43 (80)	.69	.42
Female	10 (24)	11 (28)	11 (20)		
Age (years), mean (SD)	40.4 (10.4)	37.6 (8.0)	42.7 (10.4)	.30	.01
Marital status, n (%)					
Single	11 (26)	15 (38)	12 (22)	.73	.06
Married	24 (57)	23 (58)	31 (57)		
Other	6 (14)	2 (5)	11 (20)		
Missing	1 (2)	0 (0)	0 (0)		
Monthly personal income (HK$), n (%)					
<$10,000	10 (24)	8 (20)	14 (26)	.07	.13
$10,000-$19,999	17 (41)	16 (40)	28 (52)		
$20,000-$29,999	12 (29)	6 (15)	5 (9)		
≥$30,000	2 (5)	8 (20)	6 (11)		
Missing	1 (2)	2 (5)	1 (2.0)		
FTND, [b] n (%)					
Mild	14 (33)	12 (30)	14 (26)	.61	.78
Moderate	14 (33)	18 (45)	23 (43)		
Severe	14 (33)	10 (25)	17 (32)		
Any quit attempt before intake, n (%)	35 (83)	31 (78)	36 (67)	.07	.25
Daily cigarette consumption, mean (SD)	14.5 (6.3)	15.3 (6.5)	17.1 (7.3)		

[a] Chi-square test for categorical variables; *t* test for continuous variables.

[b] FTND: Fagerstrom Test for Nicotine Dependence (1-3=mild, 4-5=moderate, 6-10=severe).

Table 2. Treatment condition and quitting characteristics of participants at baseline (N=136).

Treatment condition and quitting characteristics at baseline	Group A: WhatsApp n=42	Group B: Facebook n=40	Group C: control n=54	*P* [a]	
				A vs C	B vs C
Had been prescribed NRT, n (%)	38 (91)	39 (98)	51 (94)	.43	.56
Had been prescribed varenicline, n (%)	2 (5)	0 (0)	3 (6)	.71	.22
Frequency of smoking urge in past week, n (%)				.11	.09
Never	5 (12)	12 (30)	8 (15)		
Occasionally	20 (48)	13 (33)	13 (24)		
1-2 times per day	14 (33)	9 (23)	25 (46)		
≥3 times per day	3 (7)	6 (15)	8 (15)		
Intensity of smoking urge in past 24 hours, n (%)				.40	.10
No urge	14 (33)	19 (48)	16 (30)		
Slight	18 (43)	17 (43)	27 (50)		
Mild	8 (19)	3 (8)	11 (20)		
Moderate/Severe	2 (5)	1 (3)	0 (0)		
Frequency of thinking of the feeling of enjoying smoking, n (%)				.62	.10
Never	7 (17)	9 (23)	7 (13)		
Seldom	18 (43)	20 (50)	19 (35)		
Sometimes	15 (36)	9 (23)	25 (46)		
Often	1 (2)	2 (5)	3 (6)		
Very often	1 (2)	0 (0)	0 (0)		
Minnesota Nicotine Withdrawal Scale-Chinese, [b] mean (SD)					
Negative affect	0.46 (0.49)	0.25 (0.53)	0.48 (0.69)	.44	.02
Insomnia	0.39 (0.73)	0.60 (0.88)	0.64 (0.72)	.01	.43
12-item Smoking Self-Efficacy (SEQ-12), [c] mean (SD)					
Internal stimuli	3.78 (0.94)	3.92 (1.02)	3.89 (0.78)	.75	.56
External stimuli	4.03 (0.89)	4.11 (0.91)	4.11 (0.72)	.89	.43
Days of abstinence at baseline, mean (SD)	46.8 (16.3)	50.4 (10.5)	46.3 (16.9)	.64	.44
Days of abstinence at baseline (category), n (%)				.49	.26
≤7 days	0 (0)	0 (0)	3 (6)		
8-14 days	3 (7)	0 (0)	2 (4)		
15-28 days	5 (12)	2 (5)	4 (7)		
≥28 days	34 (81)	38 (95)	45 (83)		

[a] Chi-square test for categorical variables; Mann-Whitney *U* test for continuous variables.

[b] Greater values indicate stronger self-rated withdrawal symptoms.

[c] Greater values indicate higher self-efficacy.

Table 3. Relapsed, lapsed, smoked in the past 7 days, and validated abstinence at 2- and 6-month follow-ups.

Quitting outcomes	Group, n (%)			Unadjusted OR (95% CI)		Adjusted OR[a] (95% CI)	
	Group A (n=42)	Group B (n=40)	Group C (n=54)	A vs C	B vs C	A vs C	B vs C
2-month follow-up							
Relapse [b]							
ITT	7 (17)	12 (30)	23 (43)	0.27 (0.10, 0.71)[e]	0.58 (0.24, 1.37)	0.26 (0.09, 0.74)[f]	0.47 (0.18, 1.25)
LOCF	3 (7)	10 (25)	13 (24)	0.27 (0.07, 1.03)	1.17 (0.45, 3.05)	0.22 (0.05, 0.95)[f]	1.06 (0.35, 3.20)
Complete case	2/37 (5)	10/38 (26)	13/43 (30)	0.15 (0.03, 0.71)[f]	0.92 (0.35, 2.47)	0.17 (0.04, 0.77)[f]	0.76 (0.24, 2.39)
Lapse[c]	16 (38)	15 (38)	27 (50)	0.62 (0.27, 1.40)	0.60 (0.26, 1.38)	0.65 (0.27, 1.57)	0.58 (0.23, 1.46)
Smoked in the past 7 days	7 (17)	12 (30)	24 (44)	0.25 (0.09, 0.66)[e]	0.54 (0.23, 1.27)	0.26 (0.09, 0.73)[f]	0.44 (0.17, 1.17)
Validated abstinence[d]	16 (38)	15 (38)	13 (24)	1.94 (0.80, 4.69)	1.89 (0.77, 4.63)	1.66 (0.64, 4.33)	1.64 (0.61, 4.39)
6-month follow-up							
Relapse [b]							
ITT	17 (41)	21 (53)	33 (61)	0.43 (0.19, 0.99)[f]	0.70 (0.31, 1.61)	0.35 (0.14, 0.86)[f]	0.73 (0.29, 1.83)
LOCF	11 (26)	13 (33)	20 (37)	0.60 (0.25, 1.46)	0.82 (0.35, 1.94)	0.54 (0.21, 1.40)	0.75 (0.29, 1.96)
Complete case	9/34 (27)	9/28 (32)	17/38 (45)	0.49 (0.19, 1.31)	0.59 (0.21, 1.62)	0.43 (0.15, 1.26)	0.57 (0.18, 1.79)
Lapse[c]	24 (57)	22 (55)	33 (61)	0.85 (0.37, 1.93)	0.78 (0.34, 1.78)	0.71 (0.29, 1.74)	0.81 (0.32, 2.03)
Smoked in the past 7 days	15 (36)	21 (53)	33 (61)	0.35 (0.15, 0.82)[f]	0.70 (0.31, 1.61)	0.29 (0.11, 0.72)[e]	0.68 (0.27, 1.71)
Validated abstinence[d]	11 (26)	10 (25)	8 (15)	2.04 (0.74, 5.65)	1.92 (0.68, 5.41)	1.87 (0.62, 5.63)	2.01 (0.64, 6.36)

[a] Odds ratio adjusted for age, frequency of smoking urge in past month, intensity of smoking urge in past 24 hours, and days of abstinence at baseline.

[b] Relapse was defined as smoking 5 or more cigarettes in 3 consecutive days in the past 2 and 4 months at 2- and 6-month follow-ups, respectively. Intention-to-treat (ITT) analysis assumed participants who were lost to follow-up as relapsers or smokers. Last observation carried forward (LOCF) assumed participants who were lost to follow-up as the status of previous follow-up. Complete-case analysis excluded participants who were lost to follow-up.

[c] Lapse was defined as any incidence of smoking in the past 2 and 4 months at 2- and 6-month follow-ups, respectively.

[d] Validated abstinence (by ITT) was defined as self-reported abstinence validated by tests of exhaled carbon monoxide (≤4 ppm) and salivary cotinine (≤10 ng/mL).

[e] $P<.01$.

[f] $P<.05$.

Excluding the participants in groups A and B who did not participate in the social group (n=11), the corresponding significant odds ratios comparing groups A and C at 2- and 6-month follow-ups by ITT analysis confirmed the lower odds of relapse in group A (Multimedia Appendix 3). All odds ratios comparing the relapse rate between groups B and C were inconsistent and insignificant.

There was no significant difference in the lapse rate in all group comparisons at both follow-ups. Group A had a lower prevalence of 7-day self-reported smoking at 2-month (17%, 7/42 vs 44%, 24/54; OR 0.25, 95% CI 0.09-0.66, P=.005; power=83.7%) and 6-month (36%, 15/36 vs 61%, 33/61; OR 0.35, 95% CI 0.15-0.82, P=.02; power=67.6%) follow-ups. No significant difference between groups B and C was found. The 2-month participation rates of biochemical validation for the groups A, B, and C were 77% (20/26), 64% (16/25), and 32% (13/41), respectively. The corresponding figure at 6-month follow-up was 41% (11/27), 53% (10/19), and 33% (10/30).

The prevalence of biochemically validated abstinence at both follow-ups in groups A and B were slightly higher than group C, but the difference was insignificant.

Post Characteristics in the Social Groups

A total of 7 WhatsApp and 6 Facebook social groups were formed for groups A and B, respectively (Figure 2). Those who could not be added to the social groups (n=3) or left the social groups early (n=8) were considered "dropouts." The number of dropouts in groups A and B was 7 of 42 (17%) and 4 of 40 (10%), respectively. The mean number of posts in the WhatsApp and Facebook social groups was 55.0 (SD 50.7) and 21.0 (SD 34.4), respectively. The WhatsApp social groups had more moderators' posts (median 60, IQR 25 vs median 31.5, IQR 7; *P*=.05) and participants' posts than Facebook (median 35, IQR 50 vs median 6, IQR 9; *P*=.07), but they did not meet statistical significance. In all, 23 of 42 (54.8%) WhatsApp participants posted 1 to 9 times in the social group (median 3, IQR 7), whereas approximately half (58%, 23/40) of the Facebook participants did not post any (median 0, IQR 2.3). One WhatsApp social group had only 5 participants' posts because it had only 3 participants and 2 of them dropped out early. One Facebook social group had no posts from participants even though the moderator had already posted 24 posts of smoking cessation reminders. The majority of posts were sharing of smoking or quitting experiences (WhatsApp: 151/384 posts, 39.3%; Facebook: 81/123 posts, 65.9%) and simple reply to the moderator's inquiry (WhatsApp: 131/384 posts, 34.1%; Facebook: 82/123 posts, 66.7%) (Table 4).

Table 4. Content analysis of the WhatsApp and Facebook social groups.

Posts characteristics	Group A: WhatsApp n=42	Group B: Facebook n=40
Number of social groups, n	7	6
Participants per social group, range	2-9	4-10
Participants who did not post anything, n (%)	5 (12)	11 (28)
Total moderators' posts, n	465	255
Moderators' posts per group, median (range)	60 (46-95)	31.5 (18-118)
Total participants' posts, n	384	123
Participants' posts per group, median (range)	35 (5-145)	6 (0-90)
Participants' posts per participant, n (%)		
0 or dropped out	10 (24)	23 (58)
1-9	23 (55)	14 (35)
10-19	7 (17)	2 (5)
≥20	2 (5)	1 (3)
Characteristic of the participants' posts, [a] n (%)		
Sharing smoking/quitting experience	151 (39.3)	81 (65.9)
Simple reply to moderator's inquiry	131 (34.1)	82 (66.7)
Self-reported lapse/relapse/maintaining abstinence	62 (16.1)	15 (12.2)
Encouragement	44 (11.5)	12 (9.8)
Reminders of quitting importance	24 (6.3)	4 (3.3)
Suggesting methods for smoking cessation	21 (5.5)	7 (5.7)
Seeking information related to smoking cessation and health	8 (2.1)	4 (3.3)
Sharing information including pictures and videos	8 (2.1)	4 (3.3)
Seeking help related to smoking cessation and health	4 (1.0)	2 (1.6)
Others	85 (22.1)	4 (3.3)

[a] For group A, total number of posts was 384; for group B, total number of posts was 123.

Figure 2. Screenshots of the WhatsApp (Left side) and Facebook (Right side) online social groups.

Smoking-Related Variables at Follow-Ups

Daily smoking urge (P=.01), urge intensity (P=.02), mean score of enjoying smoking (P=.002), and the 2 subscales of the MNWS (P<.001) showed a decline significantly from baseline to 2-month follow-up and then remained steady at 6-month follow-up in the 3 RCT groups. The change of these variables in the 3 RCT groups was not significantly different (Multimedia Appendices 4-8). The mean score of the internal and external stimuli subscale of SEQ-12 increased at 2-month follow-up (P<.001), but fell at 6-month follow-up (Multimedia Appendices 9 and 10). Group A showed a greater increase in the internal stimuli score than group C (P=.04). Multimedia Appendix 11 shows the CONSORT-EHEALTH form.

Discussion

Summary of Findings

This pilot RCT found fewer participants in group A (WhatsApp) reported relapse than the control group at 2- and 6-month follow-ups. It was consistent with higher self-reported abstinence, greater change in the internal stimuli subscale of SEQ-12, and more moderators' and participants' posts in the social groups of group A. Group B (Facebook) and the control group had a similar relapse rate and the Facebook social groups had less posts than their WhatsApp counterparts.

Interpretation

Our findings have shown that the group discussion and reminders via WhatsApp social groups for recent quitters significantly reduced relapse by 73% and 57% at 2 months and 6 months, respectively. The group discussion and reminders in the WhatsApp social groups could achieve a larger effect size of maintaining short-term abstinence than face-to-face group counseling [10,11,32]. This online platform with more interactions than the Facebook social groups increased social support and reduced relapse. Also, the reminders sent from the moderators were specially designed for relapse prevention of recent quitters. These results supported previous studies that an interactive text messaging service for preventing smoking relapse was effective and well accepted by recent quitters [14,15]. The WhatsApp social groups also enhanced tailored and immediate advice from counselors, which was beneficial for smoking cessation [21]. Further investigation of the conversation content and its association with abstinence is warranted. However, the intervention effect dissipated after the

social groups closed. It suggests a longer intervention period might extend the effectiveness.

The Facebook group achieved an insignificantly lower relapse rate than the control group, which might be attributed to the insufficient sample size and statistical power. The lower effectiveness of the Facebook social groups might also be explained by the fewer moderator and participant posts than WhatsApp. Although Facebook social groups had the same standardized moderators' reminders as WhatsApp, there was less interaction between the participants and between the participants and the moderator in the Facebook social groups than WhatsApp. Therefore, Facebook participants might receive less support in preventing smoking relapses. Such a difference might be due to the difference in usage habits between WhatsApp and Facebook users. Previous studies showed that Facebook supported dissemination of smoking cessation throughout the social network [33] and engaged a large number of smokers outside the social group [21]. However, some Facebook users used the platform with computers. They might post and reply less frequently than their WhatsApp counterparts, who could do so more frequently with a mobile phone. In addition, Facebook users might be distracted by other newsfeeds, which do not appear in WhatsApp. Our findings suggested that group discussion using online social media should consider some strategies to increase interactions, such as increasing interesting and attractive content and allowing other Facebook users to post in the social group.

The significant increase in self-efficacy in dealing with internal smoking cues such as bad mood and anxiety was found in the WhatsApp group only, which suggests that the interaction and peer support in the social groups were beneficial to manage these smoking cues. The association between the change in self-efficacy and relapse prevention warrants further exploration. However, the 3 RCT groups showed similar reduction in frequency and intensity of smoking urge, and withdrawal symptoms over the study period. This result might be due to the use of the NRT as smoking cessation treatment for the majority of participants. Also, because most had been abstinent for more than a month at baseline, their barriers of maintaining abstinence might not be smoking urge or withdrawal symptoms [34]. Future studies should test the effectiveness in unassisted quitters and quitters who had quit for a few days.

Approximately half of the participants reported smoking lapses at 6 months, which was similar to a recent exploratory study providing relapse prevention to recent quitters through text messages [14]. Our online social groups did not significantly reduce smoking lapses. It was consistent with their small change in dealing with those environmental smoking cues, as measured by the external stimuli subscale of the SEQ-12, and no interaction between group allocation and time. In turn, our intervention increased internal self-efficacy and enhanced instant feedback to the reported smoking lapses, which might effectively prevent the onset of a relapse following a lapse [35]. To improve the intervention for preventing smoking lapses, process evaluation of how it helped reduce smoking lapse is needed.

Limitations

Several limitations should be noted. Firstly, the findings may only be generalizable to recent quitters (mostly male and married) who had received prior treatment from smoking cessation clinics. Such intervention should be further tested in different groups of smokers including unassisted quitters, alcoholics, and pregnant women. Secondly, most participants had already quit for a few weeks before joining the RCT, so the present RCT did not examine if the intervention helped participants manage smoking urges and withdrawal symptoms at the early stage of quitting. Thirdly, the pilot RCT had a small sample size so that the estimates of the odds ratios were not precise. Future studies with a larger sample size are warranted to confirm and quantify the effectiveness. Fourthly, differences in the smoking quantity at intake and urge frequency at baseline were found in the 3 RCT groups, which might be due to chance or no strict concealment of group allocation. Complete allocation concealment in the enrollment staff, if feasible, should be used in future RCTs. Lastly, the biochemically validated quit rate might be biased. Some self-reported quitters were too busy or perceived the validation unnecessary for them. Also, the incentive for participating in the validation might have motivated some self-reported quitters to maintain abstinence for the reward. Future studies should use other simple validation methods or compensation schemes to increase the participation in and validity of the validation.

Conclusions

This pilot RCT has developed and provided the first preliminary evidence that group discussion and reminders via WhatsApp social group were effective to reduce smoking relapse. Clinical practice in using social networking services for relapse prevention should extend the intervention period, improve the intervention content to prevent smoking lapse and relapse, and increase interaction among participants. Future RCTs with larger sample sizes and unassisted smokers are warranted.

Acknowledgments

We are grateful to the Tung Wah Group of Hospitals Integrated Centre on Smoking Cessation (ICSC) and Tobacco Control Office of Hong Kong Department of Health for funding this project and Ms Margaret Wong, Ms Jeanny Tam, and Dr Ho Kin Sun from ICSC for their suggestions in designing the intervention and evaluation. We thank Ms Rain Wan, Mr Alex Lau, and Ms Tina Fung for their contribution as social group moderators to deliver the interventions to the exsmokers. We thank Ms Zoe Wan and Ms Christina Lam for constructing the counseling content of relapse prevention. Finally, we thank the ICSC's counselors and HKU student helpers for their assistance in participant recruitment and data collection.

Conflicts of Interest

The study was funded by Tung Wah Group of Hospitals Integrated Centre on Smoking Cessation, which was funded by Tobacco Control Office of Department of Health. Prof Tai-hing Lam is the principal investigator of the FAMILY project, which was funded by the Hong Kong Jockey Club Charities Trust. All other authors do not have connection with the tobacco, alcohol, pharmaceutical, or gaming industries, and nobody was substantially funded by one of these organizations.

Multimedia Appendix 1

Using social networking service to prevent smoking relapse: Intervention guide.

[PDF File (Adobe PDF File), 42KB - jmir_v17i10e238_app1.pdf]

Multimedia Appendix 2

Using social networking service to prevent smoking relapse: Moderator's guideline.

[PDF File (Adobe PDF File), 42KB - jmir_v17i10e238_app2.pdf]

Multimedia Appendix 3

Relapsed, lapsed, smoked in the past 7 days and validated abstinence at 2- and 6-month follow-up. (excluding 11 subjects in Group A and B who did not participate in the social groups).

[PDF File (Adobe PDF File), 38KB - jmir_v17i10e238_app3.pdf]

Multimedia Appendix 4

Percentage of subjects who had daily smoking urge in the past week.

[PDF File (Adobe PDF File), 30KB - jmir_v17i10e238_app4.pdf]

Multimedia Appendix 5

Mean score of intensity of smoking urge in the past 24 hours in quitters.

[PDF File (Adobe PDF File), 31KB - jmir_v17i10e238_app5.pdf]

Multimedia Appendix 6

Frequency of thinking of enjoying smoking in the past month.

[PDF File (Adobe PDF File), 31KB - jmir_v17i10e238_app6.pdf]

Multimedia Appendix 7

Minnesota Nicotine Withdrawal Scale (Chinese): Negative affect subscale.

[PDF File (Adobe PDF File), 34KB - jmir_v17i10e238_app7.pdf]

Multimedia Appendix 8

Minnesota Nicotine Withdrawal Scale (Chinese): Insomnia subscale.

[PDF File (Adobe PDF File), 33KB - jmir_v17i10e238_app8.pdf]

Multimedia Appendix 9

Smoking Self-efficacy: Internal Stimuli.

[PDF File (Adobe PDF File), 33KB - jmir_v17i10e238_app9.pdf]

Multimedia Appendix 10

Smoking Self-efficacy: External Stimuli.

[PDF File (Adobe PDF File), 33KB - jmir_v17i10e238_app10.pdf]

Multimedia Appendix 11

CONSORT-EHEALTH V1.6 form.

[PDF File (Adobe PDF File), 1MB - jmir_v17i10e238_app11.pdf]

References

1. World Health Organization. Parties to the WHO Framework Convention on Tobacco Control. Geneva: World Health Organization; 2012.
2. Fiore MC, Jaén CR, Baker TB, Bailey WC, Benowitz NL, Curry SJ, et al. Treating Tobacco Use and Dependence: 2008 Update. Rockville, MD: Department of Health and Human Services; 2008.
3. Segan CJ, Borland R, Greenwood KM. Can transtheoretical model measures predict relapse from the action stage of change among ex-smokers who quit after calling a quitline? Addict Behav 2006 Mar;31(3):414-428. [doi: 10.1016/j.addbeh.2005.05.023] [Medline: 16005158]
4. Bold KW, Rasheed AS, McCarthy DE, Jackson TC, Fiore MC, Baker TB. Rates and predictors of renewed quitting after relapse during a one-year follow-up among primary care patients. Ann Behav Med 2015 Feb;49(1):128-140. [doi: 10.1007/s12160-014-9627-6] [Medline: 24796541]
5. Agboola S, McNeill A, Coleman T, Leonardi Bee J. A systematic review of the effectiveness of smoking relapse prevention interventions for abstinent smokers. Addiction 2010;105:1362-1380. [doi: 10.1111/j.1360-0443.2010.02996.x]
6. West R, Zhou X. Is nicotine replacement therapy for smoking cessation effective in the “real world”? Findings from a prospective multinational cohort study. Thorax 2007;62(11):998-1002. [doi: 10.1136/thx.2007.078758]
7. Balmford J, Borland R, Hammond D, Cummings K. Adherence to and reasons for premature discontinuation from stop-smoking medications: data from the ITC Four-Country Survey. Nicotine & Tobacco Research 2011;13(2):94-102. [doi: 10.1093/ntr/ntq215]
8. Hajek P, Stead LF, West R, Jarvis M, Hartmann-Boyce J, Lancaster T. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2013;8:CD003999. [doi: 10.1002/14651858.CD003999.pub4] [Medline: 23963584]
9. Stevens VJ, Hollis JF. Preventing smoking relapse, using an individually tailored skills-training technique. J Consult Clin Psychol 1989 Jun;57(3):420-424. [Medline: 2738215]
10. Smith SS, Jorenby DE, Fiore MC, Anderson JE, Mielke MM, Beach KE, et al. Strike while the iron is hot: can stepped-care treatments resurrect relapsing smokers? J Consult Clin Psychol 2001 Jun;69(3):429-439. [Medline: 11495172]
11. Powell DR, McCann BS. The effects of a multiple treatment program and maintenance procedures on smoking cessation. Prev Med 1981 Jan;10(1):94-104. [Medline: 7232346]
12. Haug S, Schaub MP, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
13. Whittaker R, Borland R, Bullen C, Lin RB, McRobbie H, Rodgers A. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2009(4):CD006611. [doi: 10.1002/14651858.CD006611.pub2] [Medline: 19821377]
14. Snuggs S, McRobbie H, Myers K, Schmocker F, Goddard J, Hajek P. Using text messaging to prevent relapse to smoking: intervention development, practicability and client reactions. Addiction 2012;107(Suppl. 2):39-44.
15. Devries KM, Kenward MG, Free CJ. Preventing smoking relapse using text messages: analysis of data from the txt2stop trial. Nicotine & Tobacco Research 2013;15(1):77-82.
16. Census & Statistics Department (Hong Kong SAR government). Hong Kong Monthly Digest of Statistics July 2015. Hong Kong: Census & Statistics Department; 2015.
17. Maher AC, Lewis KL, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
18. Brandt CL, Dalum P, Skov-Ettrup L, Tolstrup JS. "After all--it doesn't kill you to quit smoking": an explorative analysis of the blog in a smoking cessation intervention. Scand J Public Health 2013 Nov;41(7):655-661. [doi: 10.1177/1403494813489602] [Medline: 23696257]
19. Post SD, Taylor SC, Sanders AE, Goldfarb JM, Hunt YM, Augustson EM. If you build (and moderate) it, they will come: the Smokefree Women Facebook page. J Natl Cancer Inst Monogr 2013 Dec;2013(47):206-208 [FREE Full text] [doi: 10.1093/jncimonographs/lgt019] [Medline: 24395993]
20. Ramo DE, Liu H, Prochaska JJ. A mixed-methods study of young adults' receptivity to using Facebook for smoking cessation: if you build it, will they come? Am J Health Promot 2015;29(4):e126-e135. [doi: 10.4278/ajhp.130326-QUAL-128] [Medline: 24575728]
21. Struik LL, Baskerville BN. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]

22. Selby P, van Mierlo T, Voci CS, Parent D, Cunningham AJ. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. J Med Internet Res 2010;12(3):e34 [FREE Full text] [doi: 10.2196/jmir.1340] [Medline: 20719739]
23. Tung Wah Group of Hospitals Integrated Centre of Smoking Cessation. 2015. Comprehensive smoking cessation services URL: http://www.webcitation.org/6b0yWAtVw [accessed 2015-04-14]
24. World Health Organization, Department of Mental Health and Substance Dependence. The Alcohol Use Disorders Identification Test: Guildelines for Use in Primary Care. Geneva: World Health Organization; 2001.
25. Cooke F, Bullen C, Whittaker R, McRobbie H, Chen MH, Walker N. Diagnostic accuracy of NicAlert cotinine test strips in saliva for verifying smoking status. Nicotine Tob Res 2008 Apr;10(4):607-612. [doi: 10.1080/14622200801978680] [Medline: 18418783]
26. Javors MA, Hatch JP, Lamb RJ. Cut-off levels for breath carbon monoxide as a marker for cigarette smoking. Addiction 2005 Feb;100(2):159-167. [doi: 10.1111/j.1360-0443.2004.00957.x] [Medline: 15679745]
27. Van Zundert RM, Kuntsche E, Engels RC. In the heat of the moment: alcohol consumption and smoking lapse and relapse among adolescents who have quit smoking. Drug Alcohol Depend 2012 Nov 1;126(1-2):200-205. [doi: 10.1016/j.drugalcdep.2012.05.016] [Medline: 22682658]
28. Herd N, Borland R. The natural history of quitting smoking: findings from the International Tobacco Control (ITC) Four Country Survey. Addiction 2009;104(12):2075-2087. [doi: 10.1111/j.1360-0443.2009.02731.x]
29. Yu X, Xiao D, Li B, Liu Y, Wang G, Chen J, et al. Evaluation of the Chinese versions of the Minnesota nicotine withdrawal scale and the questionnaire on smoking urges-brief. Nicotine Tob Res 2010 Jun;12(6):630-634. [doi: 10.1093/ntr/ntq063] [Medline: 20498226]
30. Leung Doris Y P, Chan Sophia S C, Lau CP, Wong V, Lam T. An evaluation of the psychometric properties of the Smoking Self-Efficacy Questionnaire (SEQ-12) among Chinese cardiac patients who smoke. Nicotine Tob Res 2008 Aug;10(8):1311-1318. [doi: 10.1080/14622200802238928] [Medline: 18686178]
31. Streiner D, Geddes J. Intention to treat analysis in clinical trials when there are missing data. Evid Based Mental Health 2001;4(3):70-71. [doi: 10.1136/ebmh.4.3.70]
32. Razavi D, Vandecasteele H, Primo C, Bodo M, Debrier F, Verbist H, et al. Maintaining abstinence from cigarette smoking: effectiveness of group counselling and factors predicting outcome. Eur J Cancer 1999 Aug;35(8):1238-1247. [Medline: 10615236]
33. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating physical activity at work: using persuasive social media for competitive step counting. In: Proceedings of the 14th International Academic MindTrek Conference. 2010 Presented at: The 14th International Academic MindTrek Conference: Envisioning Future Media Environments; September 7-10, 2010; Lisboa, Portugal.
34. Segan CJ, Borland R, Hannan A, Stillman S. The challenge of embracing a smoke-free lifestyle: a neglected area in smoking cessation programs. Health Educ Res 2008 Feb;23(1):1-9 [FREE Full text] [doi: 10.1093/her/cyl157] [Medline: 17182970]
35. Gwaltney CJ, Shiffman S, Balabanis MH, Paty JA. Dynamic self-efficacy and outcome expectancies: prediction of smoking lapse and relapse. J Abnorm Psychol 2005 Nov;114(4):661-675. [doi: 10.1037/0021-843X.114.4.661] [Medline: 16351387]

Abbreviations

AUDIT: Alcohol Use Disorders Identification Test
ICSC: Integrated Centre of Smoking Cessation
ITT: intention-to-treat
LOCF: last observation carried forward
NRT: nicotine replacement therapy
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 19.06.15; peer-reviewed by P Fassier, C Maher, A Minhas; comments to author 23.07.15; revised version received 01.09.15; accepted 23.09.15; published 22.10.15.

Please cite as:
Cheung YTD, Chan CHH, Lai CKJ, Chan WFV, Wang MP, Li HCW, Chan SSC, Lam TH
Using WhatsApp and Facebook Online Social Groups for Smoking Relapse Prevention for Recent Quitters: A Pilot Pragmatic Cluster Randomized Controlled Trial
J Med Internet Res 2015;17(10):e238
URL: http://www.jmir.org/2015/10/e238/
doi:10.2196/jmir.4829
PMID:26494159

Original Paper

Social Listening: A Content Analysis of E-Cigarette Discussions on Twitter

Heather Cole-Lewis[1], MA, MPH, PhD; Jillian Pugatch[1], MPH; Amy Sanders[1], MA; Arun Varghese[1], MEng, MPA; Susana Posada[1]; Christopher Yun[1]; Mary Schwarz[1]; Erik Augustson[2], MPH, PhD

[1]ICF International, Rockville, MD, United States

[2]National Cancer Institute, Tobacco Control Research Branch, Bethesda, MD, United States

Corresponding Author:
Erik Augustson, MPH, PhD
National Cancer Institute
Tobacco Control Research Branch
9609 Medical Center Drive
MSC 9761
Bethesda, MD, 20892
United States
Phone: 1 240 276 6774
Fax: 1 240 276 7907
Email: augustse@mail.nih.gov

Abstract

Background: Electronic cigarette (e-cigarette) use has increased in the United States, leading to active debate in the public health sphere regarding e-cigarette use and regulation. To better understand trends in e-cigarette attitudes and behaviors, public health and communication professionals can turn to the dialogue taking place on popular social media platforms such as Twitter.

Objective: The objective of this study was to conduct a content analysis to identify key conversation trends and patterns over time using historical Twitter data.

Methods: A 5-category content analysis was conducted on a random sample of tweets chosen from all publicly available tweets sent between May 1, 2013, and April 30, 2014, that matched strategic keywords related to e-cigarettes. Relevant tweets were isolated from the random sample of approximately 10,000 tweets and classified according to sentiment, user description, genre, and theme. Descriptive analyses including univariate and bivariate associations, as well as correlation analyses were performed on all categories in order to identify patterns and trends.

Results: The analysis revealed an increase in e-cigarette–related tweets from May 2013 through April 2014, with tweets generally being positive; 71% of the sample tweets were classified as having a positive sentiment. The top two user categories were everyday people (65%) and individuals who are part of the e-cigarette community movement (16%). These two user groups were responsible for a majority of informational (79%) and news tweets (75%), compared to reputable news sources and foundations or organizations, which combined provided 5% of informational tweets and 12% of news tweets. Personal opinion (28%), marketing (21%), and first person e-cigarette use or intent (20%) were the three most common genres of tweets, which tended to have a positive sentiment. Marketing was the most common theme (26%), and policy and government was the second most common theme (20%), with 86% of these tweets coming from everyday people and the e-cigarette community movement combined, compared to 5% of policy and government tweets coming from government, reputable news sources, and foundations or organizations combined.

Conclusions: Everyday people and the e-cigarette community are dominant forces across several genres and themes, warranting continued monitoring to understand trends and their implications regarding public opinion, e-cigarette use, and smoking cessation. Analyzing social media trends is a meaningful way to inform public health practitioners of current sentiments regarding e-cigarettes, and this study contributes a replicable methodology.

(J Med Internet Res 2015;17(10):e243) doi:10.2196/jmir.4969

KEYWORDS

social media; Twitter; e-cigarettes; content analysis

Introduction

In recent years, electronic cigarette (e-cigarette) use has gained momentum in the United States, with 36.5% of current smokers surveyed in 2013 reporting ever using e-cigarettes, compared to 9.8% in 2010 [1]. From 2010 to 2013, e-cigarette awareness increased nearly 40%, and "ever use" of e-cigarettes increased in all demographic subpopulations except those aged 18-24 years, Hispanics, and those living in the Midwest of the United States. Results from the National Youth Tobacco Survey indicate that current e-cigarette use among high-school students tripled from the previous year, which marks the first time that current e-cigarette use has surpassed current use of every other tobacco product [2]. Data from the recent cycle of the Health Information National Trends Survey (HINTS 4 Cycle 2) indicate that among people aware of e-cigarettes, 51% believe e-cigarettes are less harmful than conventional cigarettes [3]. The growing popularity of e-cigarettes within the United States and worldwide has resulted in a surge of research spanning topics such as harm reduction, use patterns (cessation vs dual use), health effects, environmental effects, marketing, and product design. To date, data providing guidance to public health decision makers is still being collected and there is an active debate within the public health sphere regarding e-cigarette use and regulation [4]. Current findings and opinions range considerably. While some believe e-cigarettes are helping to end the global morbidity and mortality associated with the use of combustible tobacco, others believe that e-cigarettes threaten to prolong or worsen the tobacco epidemic [4-8]. Although these stances are part of the ongoing debate, there is general support for the regulation of e-cigarettes. Experts within the United States have urged the US Food and Drug Administration (FDA) to take the necessary actions to bring e-cigarettes under their regulatory authority [9], and the FDA has signaled their intention to do so [10]. However, data collection that fully addresses the public health impact of e-cigarettes and the implementation of e-cigarette regulation may take some time. In the meantime, e-cigarette sales are climbing rapidly. E-cigarette retail sales in 2013 were approximately US $2.5 billion worldwide. With this current trajectory, it is projected that retail sales in 2017 will top US $10 billion [11,12].

Given that e-cigarettes are still relatively new and the opinions toward them are often divergent, there is increasing dialogue surrounding e-cigarettes on social media. As we aim to understand the health effects of e-cigarettes, we must also attempt to discern the core voices, message frames, and sentiment surrounding e-cigarette discussions. Understanding these conversations allows public health and communication professionals to identify trends in attitudes and behaviors and to develop strategies to disseminate factual information and create culturally relevant cessation interventions for nicotine products, including traditional cigarettes and e-cigarettes.

Analysis of Twitter data has become an active research area, offering insight for the behavioral and social sciences and providing access to demographic groups that are often underrepresented in research, such as minorities. According to the Pew Research Center [13], Twitter usage is particularly popular among "younger adults, urban dwellers, and non-whites." Twitter offers greater representation of minority groups with 25% of Latinos, 27% of blacks, and 21% of whites using the social media platform [14]. In addition, 31% of 18-29-year olds use Twitter, as compared to 18% of all adult Internet users [13]. The microblog format acts as a social support system for sharing information, ideas, and beliefs that can be captured in real-time, offering insight that survey analyses might miss or take an extended period to collect [15].

In the case of tobacco use and cessation, examination of social media data can continue to uncover trends in knowledge, attitudes, and behavior; identify marketing strategies; inform public health and public policy; and pave the way for interventions delivered via social media [16-22]. Use of social media information to detect public health trends in this way has been referred to as "infoveillance" or "infodemiology" [23-26], and as digital epidemiology or digital disease detection [27].

Given the depth of data, the breadth of its audience, and its ability to capture real-time trends, this study focuses exclusively on understanding snapshots of dialogue surrounding e-cigarettes captured on Twitter. The objective of this analysis is to conduct a content analysis to identify key conversation trends and patterns over time using historical Twitter data.

Methods

To conduct this analysis, we used strategic keywords to collect historical tweets potentially related to e-cigarettes from May 1, 2013, to May 1, 2014. Keywords were selected using an iterative process with incremental addition and subtraction of words. A preliminary search, using words like e-cigarette and vapor was conducted, followed by refinement to remove terms capturing tweets that were not relevant. Addition of words was heavily influenced by a list of previously published keywords [22]. We ran several sample searches using Radian 6, social media insight software, and informally tested the performance of the searches using frequencies, correlation tables, and descriptive visualizations such as bar graphs. Product names were omitted from search keywords because we were unable to identify an exhaustive list of every e-cigarette product name and did not want to bias results to specific brands. The final list of strategic keywords is provided in Multimedia Appendix 1.

Data were provided by Gnip, a company with full access to the Twitter Firehose (entire stream of Twitter data) supplying historical tweets not available through the Twitter application program interface (API). Search results garnered 3.7 million potentially relevant tweets. Gnip data utilized for the purposes of this study include time, date, user profile link, tweet content, and tweet link. To facilitate user-friendly evaluation of the tweets among 6 analysts, a database and Web form were developed that prepopulated each tweet along with the coding categories (see Multimedia Appendices 2 and 3). Each tweet included a link to the Twitter post and the user profile. In instances where the tweet link had been removed, analysts used a historic Web cache to capture the information. Analysts visited full webpage and user profiles at their discretion when sufficient information about the tweet could not be obtained from the extraction. We did not track the number of webpage and user profile clicks made by analysts, but anecdotally, analysts clicked

these links for approximately half of all coded tweets. All analysts were college educated and participated in training sessions to familiarize themselves with e-cigarette topics and tweet analysis techniques with the lead author prior to analyzing the tweets.

Manual content analysis was used to categorize tweets according to a coding category list developed through previous literature and adapted for the purposes of this research [22]. The content analysis consisted of two stages: (1) randomly sampling tweets from the full dataset and classifying content for e-cigarette relevance until a manageable sample of at least 10,000 relevant tweets was achieved and (2) classifying content of each relevant tweet for sentiment, user description, genre, and theme. Each stage consisted of an initial step wherein analysts coded 250 of the same tweets until an acceptable level of interrater agreement was reached. Interrater reliability was determined using the Fleiss kappa and a score of at least .64 was obtained for each category in Stages 1 and 2, indicating substantial or good agreement [28,29]. Classification for relevance during Stage 1 excluded tweets that met any of the following specifications: "retweets" that offered no additional information from the person posting the tweet, original tweets that were part of a conversation and require greater context to be interpreted, or duplicated tweets from a user account that had since been suspended or was primarily being used for spam or unwanted solicitations. Spam was identified according to guidelines outlined by Twitter and included consideration of factors such as updates that are mostly links and not personal updates, duplicate content posted over multiple accounts, and content that consists of unrelated hashtags of popular topics [30,31].

Classification for sentiment, user description, genre, and theme in Stage 2 was conducted according to a codebook developed for the classification of tweets that builds on previous research [17,22] and the focus of this analysis (see Multimedia Appendix 3). Sample tweets for each content category are included in Multimedia Appendix 4).

Sentiment refers to whether the stance in the tweet is positive, neutral, or negative toward e-cigarettes and users of e-cigarettes (Table 1). Stance toward e-cigarettes and users of e-cigarettes were the only considerations for sentiment, and sentiment toward any other topic or concept was disregarded for the purposes of this research.

Table 1. Content categories for sentiment.

Category	Definition
Positive	Tweets that are in favor of e-cigarettes, related products, and use
Neutral	Tweets not strong in either direction for or against e-cigarettes
Negative	Tweets with that are against e-cigarettes

User description characterizes the sender of the tweet based on information gleaned from the user profile (eg, e-cigarette company, everyday user of Twitter, reputable news source; see Table 2). In particular, the user profile category of "e-cig community movement" was added specifically to represent those people who appear to be strong advocates for e-cigarettes but have no identified affiliation with marketing or e-cigarette companies. This category is distinct from the "everyday person" category in that the great majority of the user's timeline is devoted to e-cigarette advocacy and information, with little mention of the activities of day-to-day life characteristic of those who use Twitter for personal purposes.

Table 2. Content categories for user description.

Category	Definition
Celebrity	Famous people in pop culture, people that are Internet famous, people that have accounts verified by Twitter
Government	National Institutes of Health, CDC, political figures, etc
Foundation/organization	Reputable organizations such as American Heart Association
Reputable news source	New sources such as New York Times, Washington Post, Wall Street Journal, Associated Press, etc
Everyday person	Twitter account with a reasonable amount of posts, followers, and following a reasonable amount of people; timelines span a variety of topics that are not primarily e-cigarette–related
E-cigarette community movement	Groups or people whose timelines are primarily devoted to e-cigarette conversation (eg, Women Who Vape, The Vape Club, John Doe with entire timeline of e-cigarette tweets)
Retailer	Outlets that sell e-cigarettes (online or physical)
Tobacco company	Companies that manufacture e-cigarettes (eg, blu, Apollo, Njoy)
Bot/hacked	Accounts that appear to be fake/computerized that are primarily promoting e-cigarette products (or other products); most accounts are disguised to appear as "everyday person"

Genre represents the format of the tweet (eg, news or update, first person experience, marketing; see Table 3). *Theme*, the most granular level of classification, refers to the topical domain of the content in the tweet (eg, cessation, health and safety, craving; see Table 4).

Table 3. Content categories for genre.

Category	Definition
News/update	Update about a current event from a reputable news source, or post from user about relevant news from news source
Information	Factoid or resource, can be a personal blog or forum, or link to product review (posted by everyday person or e-cigarette community movement)
First person e-cigarette use or intent	Reports personal use of, intent, or interest to use e-cigarettes
Second/third person experience	Reports someone else's use of e-cigarette
Personal opinion	Personal opinion related to e-cigarettes
Marketing	Activities involved in the transfer of goods from the producer or seller to the consumer or buyer, eg, sales of e-cigarette products or accessories, job announcements, review of products posted by e-cigarette company/retailer

Table 4. Content categories for theme.

Category	Definition
Cessation	Mention of using e-cigarettes to quit smoking cigarettes or other non-e-cigarette tobacco products
Health and safety	Direct or indirect reference to health consequences of e-cigarette use
Underage usage	E-cig use by minors, especially high-school age or under
Craving	Desire to use e-cigarettes; eg "Stressful day. Time for my #vapepen"
Other substances	E-cigarettes mentioned in association with other addictive substances (eg, alcohol, caffeine)
Illicit substance use in e-cigarettes	Mention of using e-cigarettes for anything other than nicotine (eg, marijuana)
Policy or government	Mention of government or policy in relation to e-cigarettes including regulation, deeming, bans, and restrictions
Parental use of e-cigarettes	Tweet mentioning use of e-cigarettes by parents of the poster or parents of a person mentioned in the tweet
Advertisement/ promotion	Ads for e-cigarettes, giveaways, samples, sales, direct links to sellers' websites, word-of-mouth, and reviews
Flavors	Tweet discussing e-cigarette flavors (generic or mixed, including menthol)

Sentiment and user description are mutually exclusive categories—meaning that only one choice could be made per category, while genre and theme are not—meaning that more than one choice could be made per category. All categories were mandatory with the exception of theme, given the granularity of the content and because every topic could not be realistically represented. Additionally, during Stage 2, analysts documented media links included in each tweet (eg, image, video, location, website).

After the content analysis was complete, descriptive statistical analyses were performed on the data sample, including one-way frequencies for each category; two-way cross tabulations for categories, temporal trends, and media type, in addition to the chi-square test for intercategory statistical association (using Fisher's exact test for cell counts 5); and intercategory correlation analysis based on Cramer's V coefficient (representing each category option as a binary variable). Both the chi-square tests and correlation analyses with Cramer's V provide a statistically sound assessment of the significance and strength of the relationships between various categories. SAS version 9.3 was used for all analyses. The goal of the current analysis was to identify patterns and trends in the sample of tweets related to the overarching content categories: sentiment, user description, genre, and theme.

General trends are reported for the entire sample of coded tweets; only statistically significant trends are discussed for each category ($P<.05$). Additionally, intercategory trends are reported, once again discussing only the principal statistically significant findings ($P<.05$) of interest based on bivariate associations and intercategory correlation assessments.

Results

Sample Description

A total of 17,098 tweets were coded during Stage 1, of which 10,128 (59.23%) were found to be relevant and interpretable. The range of interrater reliability was .64-.70 and is reported in Table 5. Of the excluded tweets, 2384 were found to be entirely non-relevant, whereas the remainder were retweets with no additional context, conversations without context, or duplicated tweets from a user account that had since been suspended or was primarily being used for spam or unwanted solicitations. For the remainder of this discussion, the final sample consisted of the 10,128 relevant tweets.

Table 5. Interrater reliability scores for manual annotation of tweet categories.

Category	Interrater reliability
Relevance[a]	.70
Sentiment[b]	.65
User description[b]	.66
Genre	.64
Theme	.65

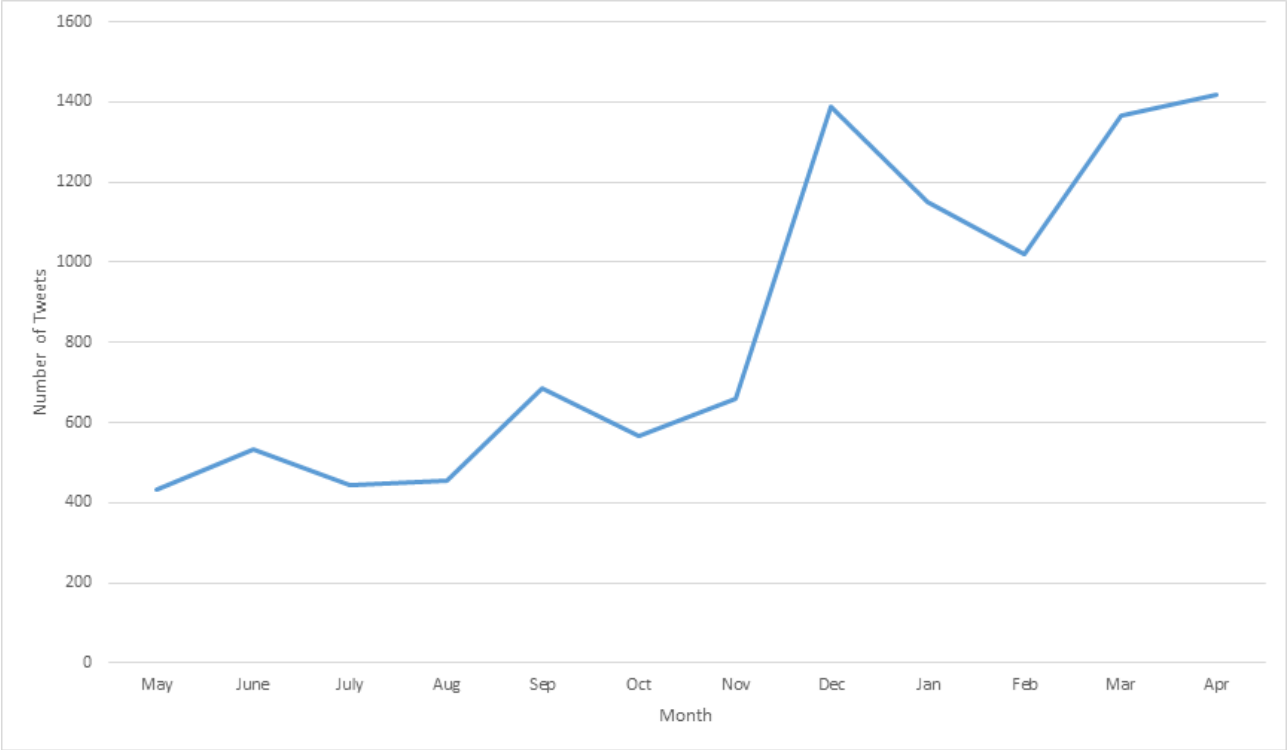
[a]Binary version of this category was created in addition to multiclass version for the purposes of the analysis.

[b]Categories were mutually exclusive and thus analyzed as multiclass.

Between May 2013 and November 2013, each month contributed 4.29-6.53% of the tweets in the overall sample; however, there is a clear increase in the number of relevant e-cigarette tweets in December 2013. The number of tweets in December 2013 (n=1388) is more than twice the number of tweets that occurred in November 2013 (n=631; see Figure 1). Months between December 2013 and April 2014 each represent 10.09% to 14.01% of the total tweets in the sample, which represents a bulk of the tweets that occurred during the observation period (see Figure 1).

Almost half of the tweets (48.00%) included links that were functional at the time of the content analysis. Tweets with images accounted for 8.30% of the sample.

Figure 1. Frequency of e-cigarette tweets by month from May 2013 to April 2014.



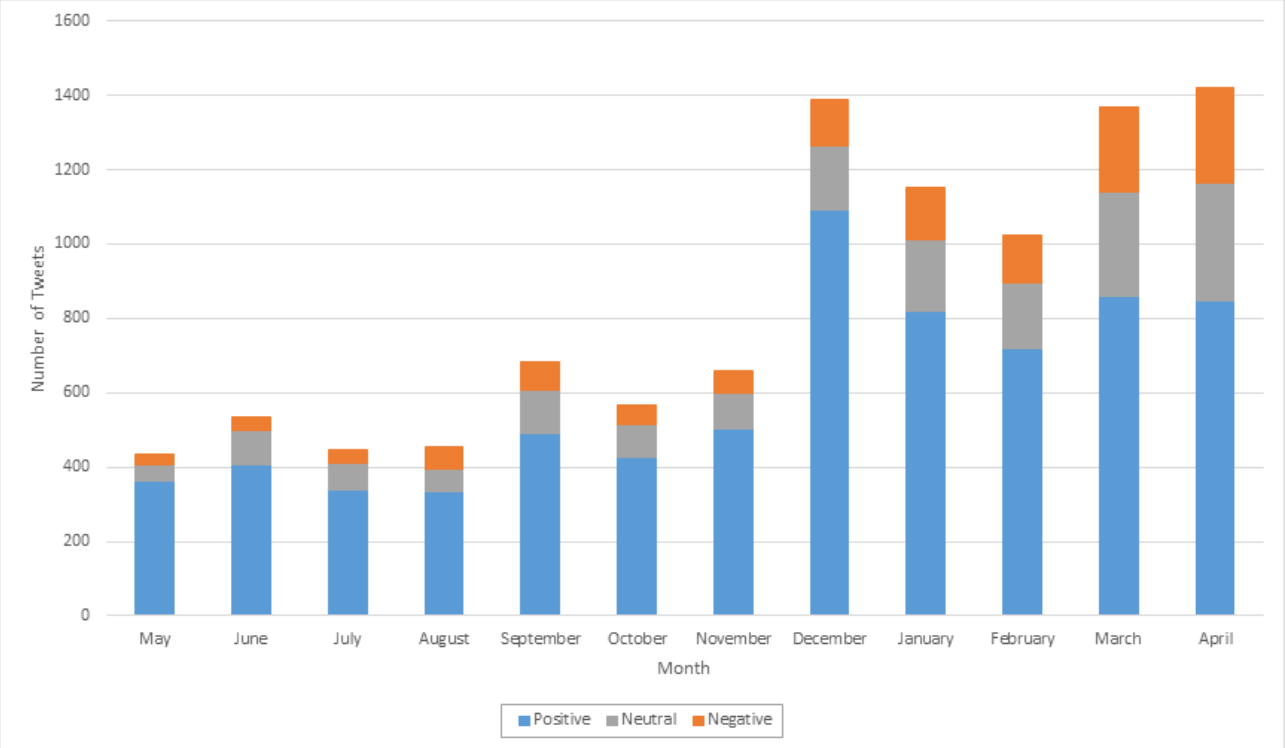
Sentiment

As indicated by Table 6, tweets deemed positive in sentiment accounted for a majority of the sample (71.11%). The absolute number of positive tweets was highest in December 2013, but May 2013 had the highest percentage of positive sentiment tweets (Figure 2). There was a steady decline in positive sentiment from December 2013 through April 2014, during which the percentage of negative and neutral tweets rose. This resulted in April 2014 having the highest percentage of negative (17.90%) and neutral tweets (22.48%).

Table 6. Tweet distribution by sentiment (N=10,128).

Sentiment	N (%)
Positive	7202 (71.11)
Neutral	1699 (16.78)
Negative	1227 (12.11)

Figure 2. Absolute number of tweets by sentiment and month from May 2013 to April 2014.



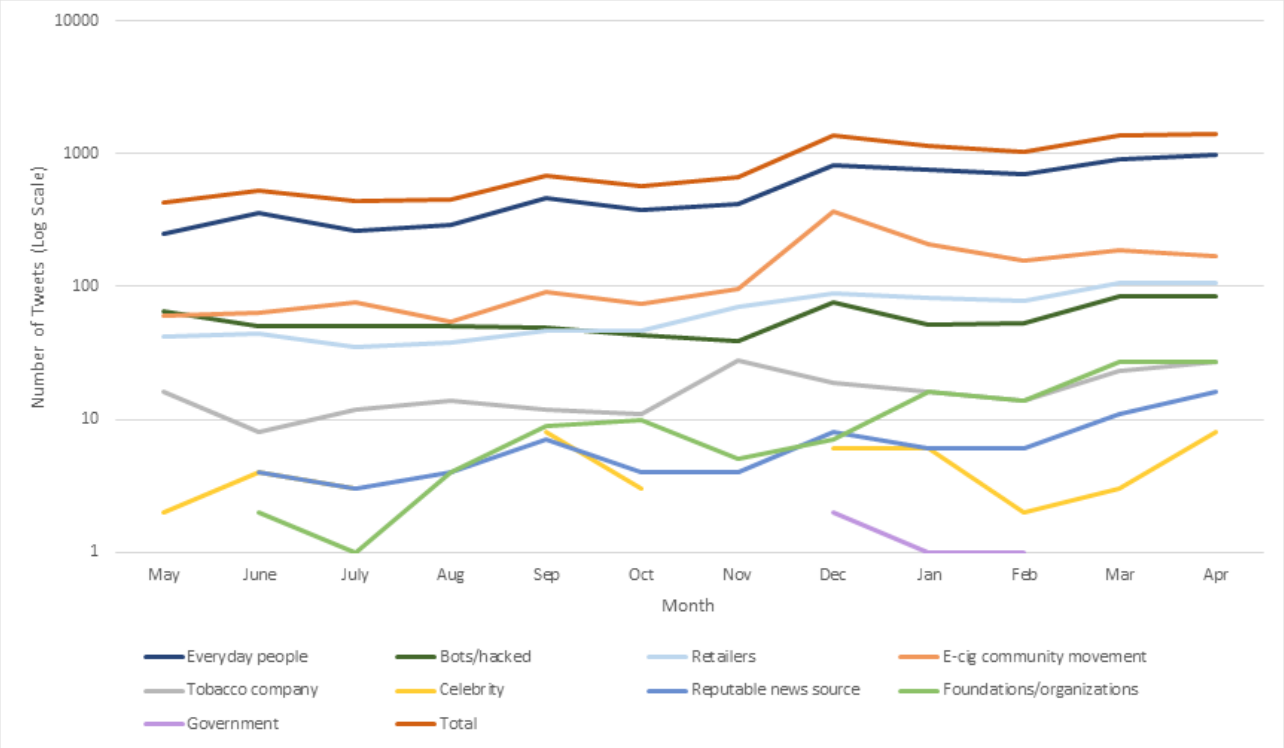
User Description

A majority of the sample consisted of tweets that originated from users identified by analysts as everyday people (64.99%), with the second largest population being the e-cigarette community (15.92%) (see Table 7). Tweets originating from government, celebrity, and reputable news sources user accounts each represented less than 1% of the sample. November 2013 saw the highest percentage of tweets from retailers (10.59%) and tobacco companies (4.24%), while the proportion of tweets from e-cigarette community movement users peaked in December 2013 (26.72%) (see Figure 3).

Table 7. Tweet distribution by user description (N=10,128).

User description	N (%)
Celebrity	45 (0.44)
Government	8 (0.08)
Foundations/organization	122 (1.20)
Reputable news source	73 (0.72)
Everyday person	6582 (64.99)
E-cigarette community movement	1612 (15.92)
Retailer	787 (7.77)
Tobacco company	200 (1.97)
Bot/hacked	699 (6.90)

Figure 3. Number of tweets by month and user from May 2013 to April 2014.



Genre

The three most common tweet genres were personal opinion-oriented tweets, marketing-related tweets, and personal experience-related tweets (see Table 8). The monthly volume of personal opinion-related tweets more than doubled during the analysis period, with a marked increase occurring from October 2013 to a peak in March 2014, though the percentage of these tweets was highest in February 2014 (see Table 9). Marketing-related tweets saw a fairly steady decline between May 2013 (29.95%) and April 2014 (18.46%) although the absolute number of these tweets doubled in that period. Personal opinion-related tweets accounted for more than one-fifth of each month's tweets, with a spike in volume occurring in December 2013, which accounted for nearly half of that months' e-cigarette related tweets (44.52%). The percentage of news-related tweets increased from 1.84% to 15.22% from May 2013 to April 2014, which represents a 27-fold increase in the volume of news-related tweets during that time. Marketing, news, and information-related tweets have much higher rates of website links (60.12%, 88.51%, and 70.40%) than average (35.42%).

Table 8. Tweet distribution by genre (N=10,128).

Genre	N (%)
News/update	828 (8.18)
Information	1459 (14.41)
First person e-cigarette use or intent	2056 (20.30)
Second/Third person experience	797 (7.87)
Personal opinion	2850 (28.14)
Marketing	2142 (21.15)

Table 9. Tweet genre distribution by month (N=10,128).

Genre	N (%)												
	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	Total
Personal experience	102 (23.50)	122 (22.90)	109 (24.49)	106 (23.35)	131 (19.12)	134 (23.63)	152 (23.00)	196 (14.12)	230 (19.97)	248 (24.27)	270 (19.74)	256 (18.04)	2056
Marketing	130 (29.95)	140 (26.27)	128 (28.76)	115 (25.33)	157 (22.92)	130 (22.93)	167 (25.26)	239 (17.22)	225 (19.53)	174 (17.03)	272 (19.88)	262 (18.46)	2139
Personal opinion	95 (21.89)	112 (21.01)	97 (21.80)	105 (23.13)	191 (27.88)	137 (24.16)	176 (26.63)	618 (44.52)	363 (31.51)	285 (27.89)	349 (25.51)	321 (22.62)	2849
Second person	34 (7.83)	52 (9.76)	34 (7.64)	45 (9.91)	50 (7.30)	37 (6.53)	48 (7.26)	82 (5.91)	96 (8.33)	93 (9.10)	112 (8.19)	114 (8.00)	797
Information	65 (15.00)	74 (13.88)	51 (11.46)	62 (13.66)	99 (14.45)	93 (16.40)	82 (12.41)	165 (11.89)	150 (13.02)	145 (14.19)	221 (16.15)	249 (17.55)	1456
News	8 (1.84)	33 (6.19)	26 (5.84)	21 (4.63)	57 (8.32)	36 (6.35)	35 (5.30)	87 (6.27)	88 (7.64)	77 (7.53)	143 (10.45)	216 (15.22)	827
Total	434	533	445	454	685	567	661	1388	1152	1022	1368	1419	10,128

Theme

Table 10 describes the overall themes present in the dataset. During the coding process, tweet theme was not a mutually exclusive category, and this resulted in 26.35% of tweets in the sample having more than one theme. For tweets with one theme, advertising and promotions-related tweets were the single largest content theme category with 19.62% occurrence in the sample, followed by policy and government–related tweets at 11.77% and health and safety–related tweets at 4.27%. Tweets coded for only the cessation theme accounted for 1.42% of the sample.

Table 10. Tweet distribution by theme.[a]

Theme	N (%)
Cessation	638 (6.30)
Health and safety	1327 (13.10)
Underage usage	423 (4.18)
Craving	394 (3.89)
Other substances	116 (1.15)
Illicit substance use in e-cigarettes	160 (1.58)
Policy/government	2042 (20.16)
Parental use of e-cigarettes	74 (0.73)
Advertisement/promotion	2663 (26.29)
Flavors	451 (4.45)

[a]Includes tweets coded with multiple themes.

Intercategory Trends

Bivariate Associations

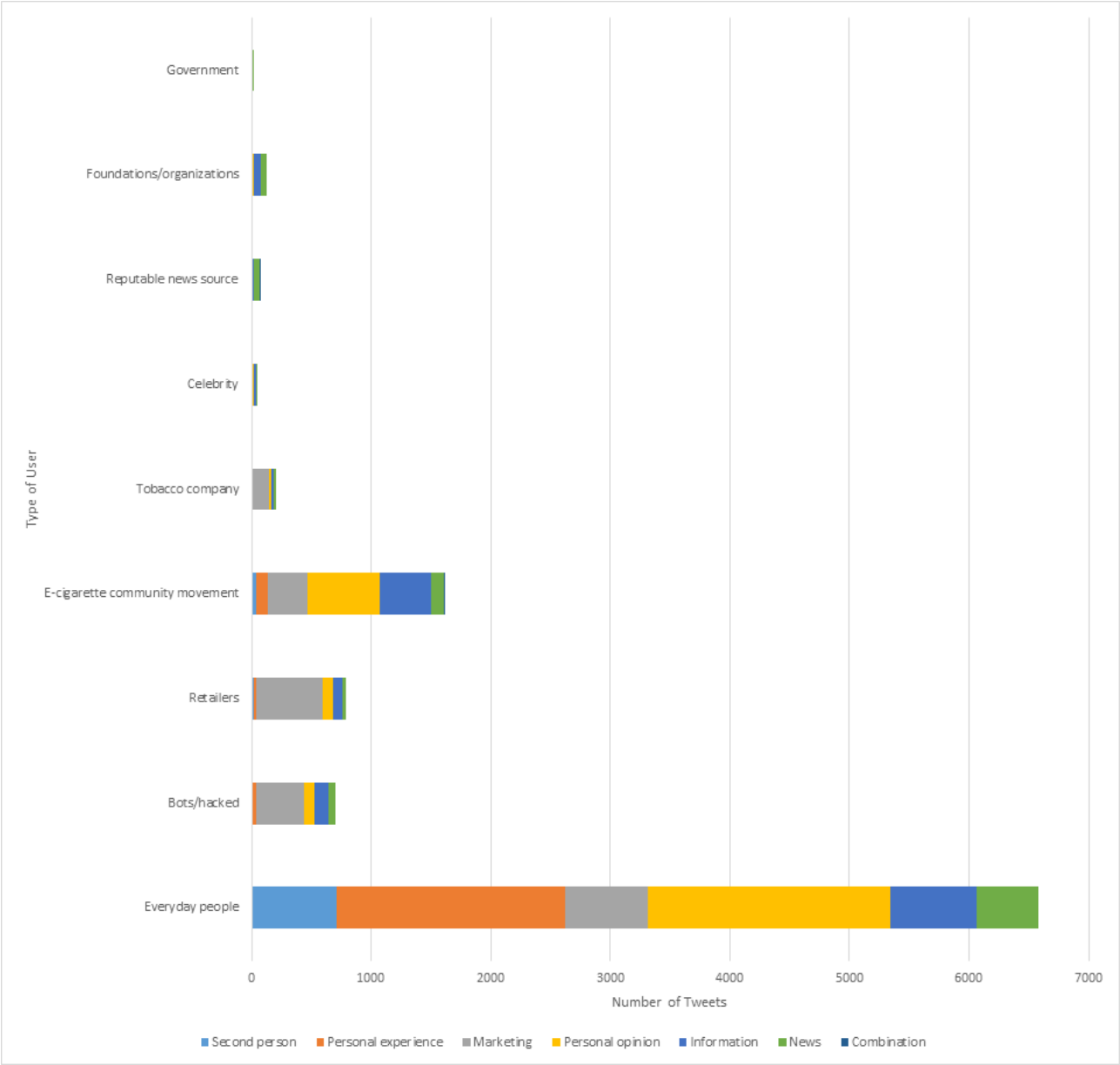
The bivariate associations reported are statistically significant ($P<.05$). Almost all marketing-related tweets were positive in sentiment (98.46%), while 88.72% of first person e-cigarette use or intent and 69.78% of personal opinion tweets were positive. Approximately half (51.65%) of informational tweets were positive and 14.22% were negative. News-related tweets were the least positive of the genres, with 19.11% of these tweets coded as positive, 53.10% neutral, and 27.81% negative.

Over 92.27% of tweets containing an image were positive in sentiment. Retailers accounted for 19.74% of tweets containing images and marketing-related tweets are twice as likely to contain an image (17.30%) compared to the average rate at which images occur in tweets (8.30%). E-cigarette community users produced 23.47% of tweets containing a link to a website. Marketing, news, and information-related tweets have much higher rates of website links (60.12%, 70.40%, and 88.51%) than the overall average (35.43%).

Nearly half (49.60%) of information tweets originated from everyday people and 29.28% were from e-cigarettes community movements. Everyday people represented 62.27% of news tweets as compared to reputable news sources accounted for 6.41% of news and 0.96% of information tweets. Foundations/organizations provided 3.98% of information tweets and 5.20% of news tweets. Also, 32.40% of marketing tweets came from everyday people compared to 26.18% from retailers

and 6.40% from tobacco companies. User-related trends in tweet genre are illustrated in Figure 4.

Figure 4. Distribution of tweet content genre for each user category.



Correlation Analysis

As an additional measure of correlation, Cramer's V statistic was computed for all categories after representing each category option as a binary variable. Multimedia Appendix 5 displays the results of the Cramer's V correlation analysis, highlighting those correlations that are moderately strong (≥.3). Results are similar to the bivariate associations discussed previously.

Discussion

Principal Findings

This content analysis revealed noteworthy trends about e-cigarette–related tweets from May 2013 through April 2014. The number of these tweets rose during the data collection period, with a peak in December 2013. Tweets were overwhelmingly positive and frequently posted by everyday people and e-cigarette community movement accounts.

The increase in e-cigarette-related tweets coincides with several e-cigarette milestones, ranging from government proposals for policies and regulations [32], major tobacco corporations introducing their own brands of e-cigarettes and buying of existing ones [33], and the e-cigarette industry's increased visibility through marketing and retail expansion. This trend may also have been driven by e-cigarette promotional activities and the subsequent increase in sales and awareness of e-cigarette products [34]. For instance, in February 2014 the company NJOY King ran an e-cigarette advertisement during the Super Bowl for a second consecutive year [35]. This televised event was viewed by over 111 million people in America, making it the most-watched television event in history at the time [36].

From May 2013 through April 2014, everyday people dominated the e-cigarette conversation on Twitter by accounting for over two-thirds of the tweets in the dataset. The e-cigarette community movement represented the second most common user type. As expected, everyday people accounted for the majority of personal opinion and personal experience tweets, though e-cigarette community movement accounts represented a sizeable proportion of personal opinion tweets as well. These two user groups accounted for 80% of information tweets, with minimal information coming from government, public health non-governmental organizations, and reputable news sources. Future research may look into the legitimacy of the information shared, its origins, and how it is shared across Twitter. This will help us understand the degree to which e-cigarette information spreads and how that might impact beliefs and opinions surrounding e-cigarettes. Everyday people also tweeted nearly a third of the marketing-related tweets, which is equivalent to the percentage of marketing tweets from retailers and tobacco companies combined. It is important to note that the volume of tweets from a particular user group does not reflect the reach or number of impressions their tweets made as this analysis did not take into account the number of followers a Twitter account has, nor the number of retweets or favorites a tweet received. Although this is a limitation of the current study, it presents the opportunity for future research to determine which Twitter voices are the "loudest" in the sense that their tweets are being seen and shared most often, and how these visible tweets influence perceptions and use of e-cigarettes.

E-cigarette community tweets spiked in December 2013, which represents a four-fold increase in tweet volume from the prior month. The cause of the sharp rise in e-cigarette community movement tweets remains unknown, but there were several e-cigarette milestones during this time. For example, in December 2013 Phillip Morris International Inc. announced its partnership with Altria Group Inc. to sell e-cigarettes [37]. A quarter of e-cigarette community tweets contained a Web link, which was considerably higher than the average. This is worth mentioning because Web links are often marketing vehicles, with 60% of the marketing tweets in this dataset containing Web links. The use of Web links among e-cigarette community users may indicate consumers' willingness to drive marketing efforts and contribute to the normalization and popularization of e-cigarettes.

Tweets originating from reputable news sources and government agencies comprised less than 1% of the sample. There continues to be debate regarding how to regulate e-cigarettes within the United States and many countries. Our analysis from Twitter suggests that the uncertainty expressed within the field of public health is not reflected in the nature of the ongoing social media dialogues. In the absence of informative dialogue from public health authorities, personal opinion and marketing content surrounding e-cigarettes have become the most common themes. This sample shows a decisive dip in tweets originating from accounts that were clearly marketers of e-cigarette products, but a large amount of marketing content continues to be posted by individuals and e-cigarette communities.

In addition to understanding who is talking on Twitter, it is necessary to dissect what is being said. Most tweets were determined to have positive sentiment indicating that Twitter dialogue skews favorably toward e-cigarettes, although the proportion of positive tweets declined during the analysis period. This trend warrants further monitoring with specific consideration for what fuels opinion over time. Furthermore, this research establishes an e-cigarette sentiment baseline that serves as a valuable starting point for public health professionals to develop campaigns and interventions. A majority of marketing-related tweets had positive sentiment, while approximately one-fifth of news-related tweets were positive. The most prevalent genre uncovered was personal opinion, followed closely by marketing to comprise nearly half of the sample. It can be expected that a platform like Twitter is conducive to sharing personal opinion; however, 32% of marketing tweets came from everyday people, while 32% came from retailers and tobacco companies combined.

Strengths and Limitations

As with any research study, our study has limitations. It must be noted that our analysis is quite specific to the topic of e-cigarettes, and thus our keyword list was limited to terms directly related to e-cigarettes. It is possible that we have overlooked conversations around topics that are socially similar to, but not exactly the same as e-cigarettes such as e-hookah. In addition, the vocabulary surrounding e-cigarettes is continuously growing and changing. This is due to the expanding range of products, brands, and vaping-related activities that people engage in. As a result, the list of keywords used in this study would need to be reconsidered and almost certainly expanded to accommodate the changing e-cigarette and vaping terminology. Furthermore, calculation of precision and recall for the search would have provided a better understanding of the validity of terms retrieved by our search. However, we are confident that our methodology is replicable, with appropriate resources and thus would allow for expansion in order to explore other emerging trends. We recommend calculation of precision and recall to refine the search and report validity using a systematic quantitative method such as that described by Stryker et al [38].

Additionally, our exclusion methodology for relevance, which included eliminating retweets without additional information and duplicate tweets from suspended accounts, may have led to an underestimation of the true prevalence of these types of tweets. However, we believe that even if our study provides a conservative estimate of the information available on Twitter in relation to e-cigarettes, the information remains useful to gain an understanding of the general trends. Future studies may be interested in utilizing less restrictive relevance criteria and using methods such as social network analysis to determine the structure of the network and how this relates to dissemination of e-cigarette attitudes and perspectives.

Twitter users do not represent the general population, and thus findings from this study must be considered in the context of people who use this specific social media platform [33]. Nonetheless, given the popularity of Twitter especially among youth, black, and Latino populations [14], information from studies such as ours provides an opportunity to access public opinions from this particular subset of the population and use

this information to form hypotheses and inform future research, as well as to supplement prior research.

Additionally, in our analysis, we did not apply an explicit weighting or correction methodology to adjust for changes in tweet volumes over time because any approach to making such an adjustment would potentially bias results. Given that most of our descriptive metrics compare fractions of tweets of a specific classification across points in time rather than absolute numbers of tweets, we believe that the comparative picture presented of the e-cig–related tweet landscape as it evolves over time is valid.

Despite a few limitations, there were many strengths of this analysis. Our study accessed data from the Twitter Firehose (ie, access to all of the daily tweets on Twitter) and utilized a large sample of tweets. Moreover, analyses were carried out over a critical period of time in the e-cigarette landscape and expanded on previously established methodologies for thematic analysis of Twitter.

An additional key strength of this work is the significant amount of time and effort spent manually building a dataset that is sampled from the Twitter Firehose (rather than the free API). The dataset of 10,128 manually coded tweets for this study is a much larger sample than previous work on Twitter and emerging tobacco products [21,22]. For example, the study presented by Huang et al included 2000 manually coded tweets, which served as the machine learning training set for over 75,000 e-cigarette-related tweets. Myslín et al's work greatly influenced our study, though it included traditional tobacco products and other emerging products (eg, hookah) in addition to e-cigarettes [22]. Myslín et al's study used over 7300 manually coded tweets for its machine learning training set, with only 4200 being relevant to tobacco, and fewer than 100 of these tweets were related to e-cigarettes.

For future research, the data from this content analysis can be used as a training dataset to build supervised machine learning algorithms. These algorithms can be used to implement automated surveillance of e-cigarette-related conversations on Twitter. This would allow more data to be analyzed with less manpower and also allow observation and analysis of Twitter trends for e-cigarette conversations over a greater period of time. This form of infoveillance lends itself to several aspects of tobacco control, including marketing regulations, underage use, cessation, and health outcomes.

Conclusion

Continuing snapshots of the social media landscape around e-cigarettes may help policymakers and public health professionals assess changing trends and inform interventions for tobacco cessation. Identifying means to integrate these types of assessments and analyses into data collected by traditional epidemiology and surveillance methods may prove especially valuable [22]. Also, this study highlights a replicable methodology and 5-category coding scheme that could be used in the future for additional topic areas.

Acknowledgments

This work was funded by the National Cancer Institute's Tobacco Control Research Branch—National Institutes of Health, National Cancer Institute HHSN261200900022C, Subcontract Number D6-ICF-1. The authors would like to thank Shinett Boggan, Alex Feith-Tiongson, Samantha Letizia, Thomas Madden, Delsie Sequiera, Nick Ngugi, and Emily Grenen for their contribution to the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tweet filter keywords.

[PDF File (Adobe PDF File), 169KB - jmir_v17i10e243_app1.pdf]

Multimedia Appendix 2

Data collection Web form.

[PDF File (Adobe PDF File), 268KB - jmir_v17i10e243_app2.pdf]

Multimedia Appendix 3

Definitions of annotation categories.

[PDF File (Adobe PDF File), 191KB - jmir_v17i10e243_app3.pdf]

Multimedia Appendix 4

Sample tweets by annotation category.

[PDF File (Adobe PDF File), 194KB - jmir_v17i10e243_app4.pdf]

Multimedia Appendix 5

Correlation matrix for content categories.

[PDF File (Adobe PDF File), 290KB - jmir_v17i10e243_app5.pdf]

References

1. King B, Patel R, Nguyen K, Dube S. Trends in awareness and use of electronic cigarettes among US adults, 2010-2013. Nicotine Tob Res 2015 Feb;17(2):219-227. [doi: 10.1093/ntr/ntu191] [Medline: 25239961]
2. Centers for Disease Control and Prevention (CDC). MMWR Morb Mortal Wkly Rep. 2015. Tobacco use among middle and high school students—United States, 2011-2014 URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6414a3.htm?s_cid=mm6414a3_e [accessed 2015-09-02] [WebCite Cache ID 6bF6LtN33]
3. Tan ASL, Bigman CA. E-cigarette awareness and perceived harmfulness: prevalence and associations with smoking-cessation outcomes. Am J Prev Med 2014 Aug;47(2):141-149 [FREE Full text] [doi: 10.1016/j.amepre.2014.02.011] [Medline: 24794422]
4. Glynn TJ. E-cigarettes and the future of tobacco control. CA Cancer J Clin 2014;64(3):164-168. [doi: 10.3322/caac.21226] [Medline: 24633865]
5. Bullen C, Howe C, Laugesen M, McRobbie H, Parag V, Williman J, et al. Electronic cigarettes for smoking cessation: a randomised controlled trial. Lancet 2013 Nov 16;382(9905):1629-1637. [doi: 10.1016/S0140-6736(13)61842-5] [Medline: 24029165]
6. Cahn Z, Siegel M. Electronic cigarettes as a harm reduction strategy for tobacco control: a step forward or a repeat of past mistakes? J Public Health Policy 2011 Feb;32(1):16-31. [doi: 10.1057/jphp.2010.41] [Medline: 21150942]
7. Polosa R, Caponnetto P, Morjaria JB, Papale G, Campagna D, Russo C. Effect of an electronic nicotine delivery device (e-Cigarette) on smoking reduction and cessation: a prospective 6-month pilot study. BMC Public Health 2011;11:786 [FREE Full text] [doi: 10.1186/1471-2458-11-786] [Medline: 21989407]
8. Callahan-Lyon P. Electronic cigarettes: human health effects. Tob Control 2014 May;23 Suppl 2:ii36-ii40 [FREE Full text] [doi: 10.1136/tobaccocontrol-2013-051470] [Medline: 24732161]
9. Gateway to addiction? A survey of popular electronic cigarettes manufacturers and targeted marketing to youth. 2014. URL: http://www.durbin.senate.gov/404 [accessed 2015-07-20] [WebCite Cache ID 6aACAA7KK]
10. U.S. Food and Drug Administration news release. Silver Spring, MD; 2014 Apr 24. FDA proposes to extend its tobacco authority to additional tobacco products, including e-cigarettes URL: http://www.fda.gov/newsevents/newsroom/pressannouncements/ucm394667.htm [accessed 2015-07-20] [WebCite Cache ID 6aACF9GwO]
11. Goodman A. Forbes. E-cigarettes are smoking hot - four ways to invest in them URL: http://www.forbes.com/sites/agoodman/2013/12/05/e-cigarettes-are-smoking-hot-4-ways-to-approach-them/ [accessed 2015-10-20] [WebCite Cache ID 6cQENsXS3]
12. Robehmed N. Forbes. E-cigarette sales surpass $1 billion as big tobacco moves in URL: http://www.forbes.com/sites/natalierobehmed/2013/09/17/e-cigarette-sales-surpass-1-billion-as-big-tobacco-moves-in/ [WebCite Cache ID 6aACIJjgF]
13. Duggan M, Smith A. Pew Research Center. 2013 Dec 13. Social Media Update 2013 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2013/Social%20Networking%202013_PDF.pdf [accessed 2015-10-20] [WebCite Cache ID 6cQEX1hj7]
14. Krogstad J. Pew Research Center. Social media preferences vary by race and ethnicity URL: http://www.pewinternet.org/files/old-media/Files/Reports/2013/Social%20Networking%202013_PDF.pdf [accessed 2015-07-20] [WebCite Cache ID 6aACPYrEZ]
15. Murphy J, Link M, Childs J, Tesfaye C, Dean E, Stern M. Social Media in Public Opinion Research: Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research. Deerfield, IL: American Association for Public Opinion Research; 2014. URL: https://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/AAPOR_Social_Media_Report_FNL.pdf [accessed 2015-07-20] [WebCite Cache ID 6aACYJYFJ]
16. Harris JK, Moreland-Russell S, Choucair B, Mansour R, Staub M, Simmons K. Tweeting for and against public health policy: response to the Chicago Department of Public Health's electronic cigarette Twitter campaign. J Med Internet Res 2014;16(10):e238 [FREE Full text] [doi: 10.2196/jmir.3622] [Medline: 25320863]
17. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. Tob Control 2012 Jul;21(4):447-449 [FREE Full text] [doi: 10.1136/tc.2010.042507] [Medline: 21730101]
18. Emery SL, Vera L, Huang J, Szczypka G. Wanna know about vaping? Patterns of message exposure, seeking and sharing information about e-cigarettes across media platforms. Tob Control 2014 Jul;23 Suppl 3:iii17-iii25 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051648] [Medline: 24935893]
19. Pechmann C, Pan L, Delucchi K, Lakon CM, Prochaska JJ. Development of a Twitter-based intervention for smoking cessation that encourages high-quality social media interactions via automessages. J Med Internet Res 2015;17(2):e50 [FREE Full text] [doi: 10.2196/jmir.3772] [Medline: 25707037]

20. Rocheleau M, Sadasivam RS, Baquis K, Stahl H, Kinney RL, Pagoto SL, et al. An observational study of social and emotional support in smoking cessation Twitter accounts: content analysis of tweets. J Med Internet Res 2015;17(1):e18 [FREE Full text] [doi: 10.2196/jmir.3768] [Medline: 25589009]
21. Huang J, Kornfield R, Szczypka G, Emery L. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014;23:30. [doi: 10.1136/tobaccocontrol-2014-051551]
22. Myslín M, Zhu S, Chapman W, Conway M. Using twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
23. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564 [FREE Full text] [doi: 10.1086/630200] [Medline: 19845471]
24. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
25. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
26. Cook S, Conrad C, Fowlkes AL, Mohebbi MH. Assessing Google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. PLoS One 2011;6(8):e23610 [FREE Full text] [doi: 10.1371/journal.pone.0023610] [Medline: 21886802]
27. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol 2015 Feb;11(2):e1003904 [FREE Full text] [doi: 10.1371/journal.pcbi.1003904] [Medline: 25664461]
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
29. Fleiss J, Levin B, Paik M. Statistical Methods for Rates and Proportions. 3rd ed. New York, NY: Wiley & Sons; 2003.
30. Twitter, Inc. 2014. The Twitter Rules URL: https://support.twitter.com/articles/18311 [accessed 2015-07-20] [WebCite Cache ID 6aACrvKQ7]
31. Twitter, Inc. 2014. Reporting Spam on Twitter URL: https://support.twitter.com/articles/64986-reporting-spam-on-twitter [accessed 2015-07-20] [WebCite Cache ID 6aACuSb0F]
32. Fehre C. A brief history of the e-cigarette.: Oxford University Press URL: http://blog.oup.com/2014/11/e-cigarette-vape-timeline/ [accessed 2015-07-20] [WebCite Cache ID 6aAD2eU9L]
33. Esterl M. The Wall Street Journal. Philip Morris to Tap E-Cigarette Market Next Year URL: http://www.wsj.com/news/articles/SB10001424052702304337404579210180309728764 [accessed 2015-07-20] [WebCite Cache ID 6aAD8ATc9]
34. Wahba P. US E-cigarette sales seen rising 24.2% per year through. 2018. URL: http://fortune.com/2014/06/10/e-cigarette-sales-rising/ [accessed 2015-07-20] [WebCite Cache ID 6aADBqg4e]
35. Gray E. Time. The super bowl ad you won't see this year URL: http://business.time.com/2014/01/23/the-superbowl-ad-you-wont-see-this-year/ [accessed 2015-07-20] [WebCite Cache ID 6aADG0x5C]
36. Nielsen. Super Bowl XLVIII Draws 111.5 Million Viewers, 25.3 Million Tweets. 2014. URL: http://www.nielsen.com/us/en/insights/news/2014/super-bowl-xlviii-draws-111-5-million-viewers-25-3-million-tweets.html [WebCite Cache ID 6aADJ8nnA]
37. Altria Group, Inc.. Altria and Philip Morris International Enter into Strategic Agreements to License and Distribute E-vapor and Other Innovative Products. 2013 Dec 20. URL: http://investor.altria.com/phoenix.zhtml?c=80855&p=irol-newsArticle&ID=1886410 [WebCite Cache ID 6aAE9cTK4]
38. Stryker JE, Wray RJ, Hornik RC, Yanovitzky I. Validation of Database Search Terms for Content Analysis: The Case of Cancer News Coverage. Journal Mass Commun Q 2006 Jun 01;83(2):413-430. [doi: 10.1177/107769900608300212]

Abbreviations

API: Application Program Interface
CDC: Centers for Disease Control and Prevention
FDA: Food and Drug Administration
HINTS: Health Information National Trends Survey
SAS: Statistical Analysis System

Edited by G Eysenbach; submitted 24.07.15; peer-reviewed by L Vera, M Conway, B Liang; comments to author 18.08.15; revised version received 22.09.15; accepted 23.09.15; published 27.10.15.

Please cite as:
Cole-Lewis H, Pugatch J, Sanders A, Varghese A, Posada S, Yun C, Schwarz M, Augustson E
Social Listening: A Content Analysis of E-Cigarette Discussions on Twitter
J Med Internet Res 2015;17(10):e243
URL: http://www.jmir.org/2015/10/e243/
doi:10.2196/jmir.4969
PMID:26508089

Original Paper

Sharing Physician Notes Through an Electronic Portal is Associated With Improved Medication Adherence: Quasi-Experimental Study

Eric Wright[1,2], PharmD, MPH; Jonathan Darer[3], MD, MPH; Xiaoqin Tang[1], PhD; Jason Thompson[1], BA; Lorraine Tusing[1], BA; Alan Fossa[4], MPH; Tom Delbanco[4], MD; Long Ngo[4*], PhD; Jan Walker[4*], RN, MBA

[1]Geisinger Health System, Center for Health Research, Danville, PA, United States

[2]Wilkes University, Department of Pharmacy Practice, Wilkes-Barre, PA, United States

[3]Geisinger Health System, Institute for Advanced Applications, Danville, PA, United States

[4]Beth Israel Deaconess Medical Center, Division of General Medicine and Primary Care, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Eric Wright, PharmD, MPH
Geisinger Health System
Center for Health Research
100 N Academy Ave
Danville, PA, 17822
United States
Phone: 1 5707146633
Fax: 1 5704087729
Email: eawright@geisinger.edu

Abstract

Background: In surveys, interviews, and focus groups, patients taking medications and offered Web portal access to their primary care physicians' (PCPs) notes report improved adherence to their regimens. However, objective confirmation has yet to be reported.

Objective: To evaluate the association between patient Internet portal access to primary care physician visit notes and medication adherence.

Methods: This study is a retrospective comparative analysis at one site of the OpenNotes quasi-experimental trial. The setting includes primary care practices at the Geisinger Health System (GHS) in Danville, Pennsylvania. Participants include patients 18 years of age or older with electronic portal access, GHS primary care physicians, and Geisinger health plan insurance, and taking at least one antihypertensive or antihyperlipidemic agent from March 2009 to June 2011. Starting in March 2010, intervention patients were invited and reminded to read their PCPs' notes. Control patients also had Web portal access throughout, but their PCPs' notes were not available. From prescription claims, adherence was assessed by using the proportion of days covered (PDC). Patients with a PDC ≥.80 were considered adherent and were compared across groups using generalized linear models.

Results: A total of 2147 patients (756 intervention participants, 35.21%; 1391 controls, 64.79%) were included in the analysis. Compared to those without access, patients invited to review notes were more adherent to antihypertensive medications—adherence rate 79.7% for intervention versus 75.3% for control group; adjusted risk ratio, 1.06 (95% CI 1.00-1.12). Adherence was similar among patient groups taking antihyperlipidemic agents—adherence rate 77.6% for intervention versus 77.3% for control group; adjusted risk ratio, 1.01 (95% CI 0.95-1.07).

Conclusions: Availability of notes following PCP visits was associated with improved adherence by patients prescribed antihypertensive, but not antihyperlipidemic, medications. As the use of fully transparent records spreads, patients invited to read their clinicians' notes may modify their behaviors in clinically valuable ways.

(J Med Internet Res 2015;17(10):e226) doi:10.2196/jmir.4872

KEYWORDS

patient portal; adherence; progress notes; hypertension; hyperlipidemia

Introduction

Advances in health information technology, including secure patient portals associated with electronic health records (EHRs), create new opportunities for providers, patients, and families to interact and share information. Moreover, patient reports suggest strongly that unprecedented access to clinical information, including the notes clinicians write, may improve adherence to medical regimens [1].

With the goal of promoting engagement and communication between patients and clinicians, primary care physicians (PCPs) participating in the OpenNotes initiative invited their patients to read progress notes through a secure patient portal. A large majority of patients in each site chose to read their notes and reported feeling more in control of their health care, being more prepared for visits, and understanding their medical conditions better [2]. Of particular interest, they reported improving their adherence to medications. More than two-thirds of patients taking medications who responded to postintervention surveys reported that access to their PCPs' notes helped them "take their medications better" [3]. A similar small pilot study in heart failure patients supports a correlation between medical record access and improved adherence, but beyond self-report, no large studies have examined whether adherence does indeed change [4].

Nonadherence to medications is a major and potentially modifiable risk factor for heavy utilization of health care resources, disease progression, and mortality [5]. For newly prescribed medications, 10 to 30% of patients never fill the original prescription, and for some medications prescribed to manage chronic illness, nonadherence rates can approach 50% at 6 months [6]. Low adherence to medications can affect both patient outcomes and health care costs, with some estimating US $100 billion in excess US health care system costs per year [7]. Improving adherence has long been a challenge for both clinicians and patients; although some interventions targeting the patient, provider, and the health delivery system have demonstrated higher adherence, they are often limited to defined diseases or patient populations, are frequently costly, and the results from different studies vary considerably [8,9]. Ideally, interventions designed to encourage adherence would take into account patient and provider preferences, would be inexpensive, would both work across and be tailored to multiple health conditions, and would be replicable across settings.

The Geisinger Health System (GHS), one of three sites in the OpenNotes inquiry, offers health insurance to about a third of the patients it manages, and among them we had the opportunity to examine to what degree patients filled and refilled prescriptions written by their PCPs. Two conditions with high frequency in this population are hypertension and hyperlipidemia. With medications prescribed for these conditions as markers, the goal of our study was to evaluate whether a reminder to read, and the ready availability of, PCP progress notes did indeed affect patient medication adherence.

Methods

Design Overview

The OpenNotes study was a multicenter, prospective quasi-experimental study of participants invited to read PCP notes through electronic portals following their visits [2]. In 2010, the intervention was initiated at three sites: Beth Israel Deaconess Medical Center in Boston, Harborview Medical Center in Seattle, and the GHS in Pennsylvania. Surveys completed by patients and physicians, both those volunteering to participate in the intervention and those declining participation, were previously reported [1,3]. For this inquiry, we conducted a quasi-experimental comparative study among patients within the GHS. The study was conducted with Institutional Review Board (IRB) review at GHS; it was determined to be exempt due to the use of deidentified data.

Setting and Participants

GHS is a fully integrated health care system that provides health services to more than 2.6 million people in a 44-county region in central and northeast Pennsylvania. In GHS's area of coverage, 29 counties are designated as rural, with 18% of the population over 65 years of age. An EHR system has been fully operational within GHS practices since 2001. Information from the EHR along with financial, health plan prescription and medical claims, patient satisfaction, and high-use third-party reference datasets are backed up on a comprehensive enterprise data warehouse every 24 hours.

Patients from GHS included those cared for by 24 PCPs in 14 clinic locations who volunteered to participate in the OpenNotes evaluation, and those cared for by 78 PCPs in the same clinics who declined to participate. There was at least one nonparticipating PCP in each clinic where a participating PCP practiced. Patients were included if they were 18 to 89 years of age, assigned a GHS PCP, had an active portal account (called MyGeisinger), and had Geisinger health plan insurance 1 year prior to the intervention (March 1, 2009-February 28, 2010 or July 1, 2009-June 30, 2010) and throughout the time period of the intervention (March 1, 2010-February 28, 2011 or July 1, 2010-June 30, 2011). Patients were included if they had a prescription claim filed for an antihypertensive or antihyperlipidemic medication during the year prior to deploying OpenNotes (baseline), and during the study year (follow-up) of the analysis.

Intervention

Intervention patients were offered access to their PCP notes via the MyGeisinger Web portal. Following signature of a note by a PCP documenting an encounter, patients received an email message sent to their personal email address notifying them of a portal message. The message included a direct link to the note section of the portal. In addition, prior to their next scheduled PCP appointment, a MyGeisinger reminder message was sent, encouraging the patient to review the previous PCP note(s) prior to the appointment. Control patients were those meeting the inclusion criteria above and were listed as a patient from the panel of nonparticipating physicians within a practice where at least one OpenNotes physician was participating. Control

patients were also MyGeisinger users, but did not have access to their PCP notes. All other portal functionality, including access to lab results, medication and problem lists, appointment information, and correspondence with providers, was available to both intervention and control patients throughout the study period. Because there were more PCPs who did not volunteer to participate, we expected to have more control than intervention patients.

Outcomes and Follow-Up

We compared adherence to two classes of medications, antihypertensive and antihyperlipidemic agents, both commonly prescribed in our primary care population and for which previous interventional studies have evaluated adherence through prescription records [10-13]. From prescription claims, adherence was assessed by using the proportion of days covered (PDC) [14]. The PDC was defined as the number of days covered, based on the medication fill date and days of supply, multiplied by the number of claims in the defined period, divided by the length of the study period (365 days). Drugs within one subclass were included in the same PDC calculation. If more than one subclass claim was identified per patient during the period of interest (eg, atenolol and hydrochlorothiazide), separate PDCs were calculated for each and then averaged together to arrive at one antihypertensive or antihyperlipidemic PDC per patient per period of interest. The value of a PDC is bounded between 0 and 1 and is often converted to percentages between 0 and 100%. Our primary outcome was adherence during the follow-up year. We used a breakpoint of ≥80% PDC for adherence, a percentage widely accepted in the literature as appropriate for designating adherent versus nonadherent patients for these classes of medications [15,16]. We evaluated rates of adherence versus nonadherence in the two time periods (baseline and follow-up).

Statistical Analysis

We calculated means and standard deviations for symmetric continuous variables, medians and interquartile ranges (IQRs) for asymmetric continuous variables, and frequencies and percentages for categorical variables. Baseline comparisons between intervention and control groups were made using a two-sample *t* test or Kruskal-Wallis test, and Pearson's chi-square tests. We used a dichotomous variable of *adherent* as the outcome variable, with a value of 1 assigned to patients having ≥80% PDC. Multivariable logistic regression was used to detect differences between groups, controlling for potential confounders which included diabetes, body mass index (BMI), and the number of primary care visits per year.

In a secondary analysis to determine whether patients changed adherence status pre- and postintervention, a four-level outcome variable was created: nonadherent to adherent, adherent to adherent, nonadherent to nonadherent, and adherent to nonadherent. A multinomial logistic regression model was used to test for differences between groups. All statistical analyses were performed using SAS software, version 9.3 (SAS Institute, Cary, North Carolina).

Results

Patients with GHS insurance cared for by 102 PCPs—24 participating and 78 nonparticipating—were available for comparison. In all, 2147 subjects were eligible and included in our analysis: 756 (35.21%) in the OpenNotes participating group (intervention) and 1391 (64.79%) in the nonparticipating group (control) (see Figure 1).

Characteristics at baseline were similar between the two groups, except for the prevalence of diabetes mellitus and an elevated BMI, both of which were slightly higher in the intervention group (see Table 1). Patients were 59 years of age on average, split about evenly between men and women, and overwhelmingly white.

Table 1. Baseline[a] characteristics for antihypertensive and antihyperlipidemic cohorts.

Variable	Antihypertensive agents			Antihyperlipidemic agents		
	Intervention (n=561)	Control (n=1008)	*P*	Intervention (n=474)	Control (n=913)	*P*
Age in years, mean (SD)	60.5 (12.9)	59.9 (12.7)	.39	60.4 (11.6)	60.6 (11.8)	.74
Sex (female), n (%)	289 (51.5)	554 (54.96)	.19	229 (48.3)	466 (51.0)	.33
Race (white), n (%)	556 (99.1)	1001 (99.31)	.67	471 (99.4)	908 (99.5)	.84
Hospitalizations in previous year, n (%)	52 (9.3)	89 (8.83)	.77	46 (9.7)	80 (8.8)	.56
Emergency department visits in previous year, n (%)	73 (13.0)	110 (10.91)	.21	61 (12.9)	90 (9.9)	.09
PCP encounters per year, median (IQR)	3.3 (2.3-5.3)	3.3 (2.0-4.7)	.09	3.3 (2.0-5.0)	3.3 (2.0-4.7)	.22
Body mass index (kg/m), median (IQR)	31.6 (28.1-35.9)	31.1 (27.5-35.4)	.08	30.7 (27.3-34.7)	30.1 (26.8-33.8)	.05
Systolic blood pressure (mmHg), mean (SD)	130.8 (12.2)	131.0 (11.5)	.74	127.9 (12.0)	128.4 (11.6)	.43
Diastolic blood pressure (mmHg), mean (SD)	76.4 (8.1)	76.2 (8.0)	.62	75.1 (7.6)	75.1 (7.4)	.91
Total cholesterol (mg/dl), mean (SD)	180.9 (34.8)	184.6 (38.1)	.07	179.8 (38.4)	182.8 (39.6)	.17
High density lipoprotein (mg/dl), mean (SD)	49.6 (14.7)	51.3 (15.1)	.04	48.8 (13.3)	50.3 (13.9)	.06
Low density lipoprotein (mg/dl), mean (SD)	100.8 (29.0)	103.3 (32.1)	.15	99.0 (31.7)	101.6 (32.8)	.16
Hemoglobin A1c (%), mean (SD)	6.8 (1.2)	6.7 (1.0)	.32	6.8 (1.1)	6.8 (1.1)	.54
Charlson Comorbidity Index score, mean (SD)	0.9 (1.1)	0.8 (1.1)	.56	0.9 (1.1)	0.8 (1.1)	.58
Coronary artery disease, n (%)	103 (18.4)	151 (14.98)	.08	94 (19.8)	177 (19.4)	.84
Heart failure, n (%)	36 (6.4)	82 (8.13)	.22	27 (5.7)	63 (6.9)	.39
Hypertension, n (%)	438 (78.1)	794 (78.77)	.75	279 (58.9)	537 (58.8)	.99
Hyperlipidemia, n (%)	321 (57.2)	586 (58.13)	.72	380 (80.2)	749 (82.0)	.40
Diabetes mellitus, n (%)	160 (28.5)	231 (22.92)	.01	137 (28.9)	224 (24.5)	.08

[a]Baseline measurements occurred on March 1, 2010, for pilot providers and July 1, 2010, for remaining participating providers.

[b]Primary care physician (PCP).

[c]Interquartile range (IQR).

[d]Average of up to three most recent readings leading up to intervention start.

[e]Most recent reading leading up to intervention start.

Of the 2147 eligible participants, 818 (38.10%) (281 intervention, 34.4%; 537 control, 65.6%) had a prescription claim for both an antihypertensive and antihyperlipidemic agent, 751 (34.98%) (280 intervention, 37.3%; 471 control, 62.7%) had a prescription claim for only an antihypertensive agent, and 569 (26.50%) (193 intervention, 33.9%; 376 control, 66.1%) had a prescription claim for only an antihyperlipidemic agent in both baseline and follow-up years. Out of 2147 patients, 1569 (73.08%) (561 intervention, 35.76%; 1008 control, 64.24%) were taking at least one antihypertensive agent, and 1387 (64.60%) (474 intervention, 34.17%; 913 control, 65.83%) were taking at least one antihyperlipidemic agent during the baseline and follow-up time periods. Average PDC ranged from 85 to 87% across the 2-year time frame for both antihypertensive and antihyperlipidemic agents (see Multimedia Appendix 1).

Patients in the intervention group were more likely to be adherent to antihypertensive medications compared to the control group (79.7% and 75.3%, respectively)—absolute risk difference, 4.4%; adjusted risk ratio, 1.06 (95% CI 1.00-1.12); *P*=.04 (see Table 2). We found no difference in adherence among those prescribed antihyperlipidemic agents (77.6% for intervention and 77.3% for control)—adjusted risk ratio, 1.01 (95% CI 0.95-1.07); *P*=.86.

Table 2. Comparison of patients' adherence[a] to treatment during baseline and follow-up periods in the intervention and control groups.

Type of agent	Intervention		Control		Unadjusted		Adjusted[b]	
	BL[c], n (%)	F-U[d], n (%)	BL, n (%)	F-U, n (%)	BL risk ratio (95% CI), *P*	F-U risk ratio (95% CI), *P*	BL risk ratio (95% CI), *P*	F-U risk ratio (95% CI), *P*
Antihypertensive (n_I[e]=561, n_C[f]=1008)	439 (78.3)	447 (79.7)	799 (79.27)	759 (75.30)	0.99 (0.94-1.04), .64	1.06 (1.00-1.12), .04	0.99 (0.94-1.05), .84	1.06 (1.00-1.12), .04
Antihyperlipidemic (n_I[e]=474, n_C[f]=913)	372 (78.5)	368 (77.6)	691 (75.7)	706 (77.3)	1.04 (0.98-1.10), .23	1.00 (0.95-1.07), .90	1.04 (0.98-1.10), .21	1.01 (0.95-1.07), .86

[a]Adherent is defined as patients with proportion of days covered (PDC) ≥.80 (percentage with PDC ≥80%).

[b]For antihypertensive agents: adjusted for diabetes, body mass index (BMI), and primary care visits per year; for antihyperlipidemic agents: adjusted for diabetes and primary care visits per year.

[c]Baseline (BL).

[d]Follow-up (F-U).

[e]Number of included patients in the intervention group (n_I).

[f]Number of included patients in the control group (n_C).

In a secondary pre/post comparative analysis of antihypertensive users, the percentages of nonadherent to adherent and persistently adherent were 11.2% (63/561) and 68.5% (384/561) in the intervention group, respectively, compared to 9.23% (93/1008) and 66.07% (666/1008) in the control group (see Table 3). Intervention patients had a 27% higher probability of having the outcome of nonadherent to adherent, versus changing from adherent to nonadherent—adjusted risk ratio, 1.27 (95% CI 1.00-1.54); *P*=.048. In a number-needed-to-treat (NNT) analysis, with an absolute difference of 2.0% improvement (11.2% vs 9.2%) from nonadherent to adherent, for every 50 patients offered access to physicians' notes through a Web portal, 1 additional patient will move from nonadherent to adherent to their antihypertensive medication. There was no significant difference between groups in movement from nonadherent to adherent among antihyperlipidemic users—adjusted risk ratio, 0.87 (95% CI 0.65-1.11); *P*=.30.

Figure 1. Patient cohort selection.

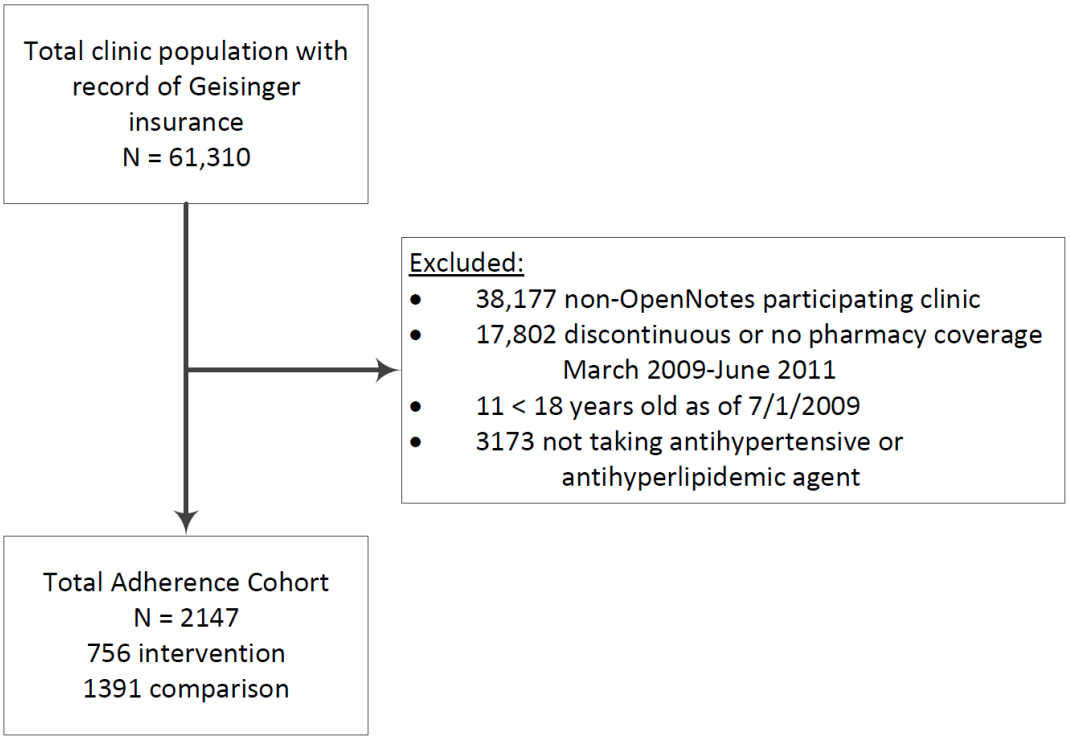


Table 3. Within-person proportional changes and risk estimates in adherence from baseline to follow-up.

Type of agent	Movement from baseline (1st listed) to follow-up (2nd listed)	Intervention, n (%)	Control, n (%)	Unadjusted RR[a] (95% CI), *P*	Adjusted[b] RR (95% CI), *P*
Antihypertensive (n_I[c]=561, n_C[d]=1008)					
	Reference (adherent[e] to nonadherent[f])	55 (9.8)	133 (13.19)	N/A[g]	N/A
	Nonadherent to adherent	63 (11.2)	93 (9.23)	1.30 (1.03-1.56), .03	1.27 (1.00-1.54), .048
	Adherent to adherent	384 (68.4)	666 (66.07)	1.05 (1.00-1.09), .05	1.05 (1.00-1.09), .04
	Nonadherent to nonadherent	59 (10.5)	116 (11.51)	1.03 (0.96-1.09), .36	1.04 (0.96-1.09), .30
Antihyperlipidemic (n_I[c]=474, n_C[d]=913)					
	Reference (adherent to nonadherent)	49 (10.3)	80 (8.8)	N/A	N/A
	Nonadherent to adherent	45 (9.5)	95 (10.4)	0.88 (0.66-1.11), .32	0.88 (0.65-1.11), .30
	Adherent to adherent	323 (68.1)	611 (66.9)	0.98 (0.93-1.02), .44	0.98 (0.92-1.02), .42
	Nonadherent to nonadherent	57 (12.0)	127 (13.9)	0.88 (0.68-1.06), .20	0.88 (0.68-1.06), .20

[a]Risk ratio (RR).

[b]For antihypertensive agents: adjusted for diabetes, body mass index (BMI), and primary care visits per year; for antihyperlipidemic agents: adjusted for diabetes and primary care visits per year.

[c]Number of included patients in the intervention group (n_I).

[d]Number of included patients in the control group (n_C).

[e]Adherent is defined as proportion of days covered (PDC) ≥80%.

[f]Nonadherent is defined as PDC <80%.

[g]Not applicable (N/A).

Discussion

This study is the first large-scale report suggesting that medication adherence to antihypertensive medications improves among patients granted access to review PCP notes through a Web portal. In individual interviews, focus groups, and surveys, patients indicated that being reminded to, and having access to, read their clinicians' notes lead them to use prescribed medications "better" [1]. This study provides evidence that a cohort of patients invited to review their PCPs' progress notes through a secure electronic portal demonstrate increased adherence to medications prescribed for hypertension, but not for hyperlipidemia.

Our results are consistent with a smaller randomized study in which 107 heart failure patients received either usual care or access to a secure online medical record that also included clinical notes [4]. Using the general adherence scale, self-reported adherence improved significantly at 12 months in the intervention group. In our study of a larger patient population, information gathered from prescription claims is consistent with that finding when evaluating patients prescribed medications for hypertension.

Why did we find change among patients with respect to antihypertensive therapy, but not with antihyperlipidemic drugs? Prior to undertaking this analysis, in a group discussion with PCPs, including two authors of this study (JD, TD), a group of clinicians hypothesized the findings would more likely demonstrate increased adherence to medications prescribed for antihypertensives than for lipid abnormalities. Reflecting on their own practices, they felt their notes frequently reflected uncertainty about indications for therapy with statins, along with concerns about side effects. In contrast, they felt notes were more definitive about the need for patients to use antihypertensives. Further, pharmacotherapy guidelines for hypertension are widely accepted by clinicians, but there is long-standing and constantly evolving debate about indications for pharmacotherapy following the measurement of lipid levels [17,18]. Do the doctors' notes convince patients to adhere more

closely to antihypertensive regimens, while reinforcing potential ambivalence in both their doctors' and their own minds when it comes to evaluating and managing lipid levels?

In a post hoc sensitivity analysis of patients prescribed both antihypertensive and antihyperlipidemic agents versus those prescribed either alone, we found a similar magnitude of adherence effect in the antihypertensive and antihyperlipidemic arms regardless of concomitant therapy (internal analysis). In essence, we noted no positive carryover of effect of using antihypertensive agents among those also using antihyperlipidemic agents. This may infer that there is indeed a differential value placed on the benefit of antihypertensive therapy versus antihyperlipidemic therapy. While our findings support our hypothesis, there remains uncertainty as to why a differential effect was found.

Another potential rationale for the difference noted between antihypertensive and antihyperlipidemic groups could relate to the very small decline in adherence from baseline year to follow-up year in either the intervention or the control groups despite evidence that adherence declines over time [19]. We suspect that the inclusion of prevalent antihypertensive and antihyperlipidemic users resulted in this more gradual decline from baseline to follow-up. A drop in adherence to these chronic medications in the first year is well established, but less is known about the adherence changes from one year to the next in populations with long-standing use of such drugs. In a study of elderly patients taking antihypertensive medications, Krousel-Wood et al reported about a 4.3% decline in the rate of adherence per year [20]. We noted a comparable adherence change in the control antihypertensive group, but not in the antihyperlipidemic control arm, suggesting that the antihypertensive control is a more reliable comparison group.

Although we demonstrated that patients who have access to their progress notes have a higher adherence rate to antihypertensive medications, in contrast to the striking patient self-reports, the magnitude of this effect was small. This may reflect limitations in study design deriving from retrospective analysis in a quasi-experimental study not designed or powered to test hypotheses about adherence to medication. We also noted very high adherence levels of our patients throughout the baseline and follow-up periods, perhaps limiting our ability to measure significant changes in adherence. The mechanism by which our patients have improved adherence is largely explained through improved engagement by reading patient notes and/or the reminder to read them. Although this mechanism seems ultimately plausible based upon this evaluation study, it does not specifically target medication nonadherence like other interventions facilitated through patient portals, such as automatic refill requests [21].

Nevertheless, the internal validity of our study is strong, due to the study design and similar comparative groups (see Table 1). We worked to overcome confounding and selection biases by using controls from offices where both participating and nonparticipating physicians cared for patients. In addition, we adjusted our analysis for known and potential confounders and found nothing measurably different from the crude results, suggesting that the comparison group contained characteristics very similar to those known and likely unknown in the intervention group. Finally, we were able to find and measure variables of interest through our electronic and administrative claims database, providing detailed capture of information among our largely nonmigratory patient population.

While the findings are consistent with our initial hypotheses, they may be confounded by factors we cannot measure at baseline. Although not randomized, since the intervention was implemented to patients on the level of the physician, it minimizes, but may not eliminate, unmeasured patient-level differences that could confound results. Differences between groups that developed during the study are more difficult to attribute to confounders unrelated to the interventions instituted as part of rolling out this project. For example, if providers increased contact and interventions with patients more in the intervention arm, this may be due to patients reading their doctors' notes, or it may be unrelated. With OpenNotes being as much a physician-level as a patient-level intervention, this change in practice may impact physician behavior in ways that could influence the outcomes measured in this study, in essence resulting in a Hawthorne effect reinforcing the change in patient behavior we anticipated would occur.

Finally, although not designed or powered to detect differences in clinical end points, medication adherence was identified as a measure of interest during the initial and subsequent phases of this project, limiting bias caused by testing of multiple hypotheses. However, while an appropriate comparator group was identified, there may be baseline differences between the groups beyond those for which we controlled in our adjusted analysis, and this could contribute to the different findings for the two classes of medications. On the other hand, the findings may also underestimate the impact of the intervention, since some patients taking either of the medications in the study group may not have read their PCPs' notes. Overall, 18% of GHS patients cared for by PCPs in the intervention group chose not to read any notes during the approximately year-long study period. While it could be expected that larger effect sizes could be reached by excluding those not having read their notes, we attempted to show the real-world effects of the intervention by including all patients, regardless of note viewing in the final analysis.

Patients reported that reading their clinicians' notes helps them with their medical regimen, and improved adherence may both improve the quality of care and decrease costs over time, thereby adding value to a system avidly seeking ways to improve care. Albeit carrying modest weight, our findings are consistent with what patients report and what we anticipated. More and wider measurements of these important components of care are urgently required, but for now, the evidence is increasing that fully transparent records can improve communication and engage patients more actively in their own care.

Acknowledgments

This study was supported in part from a grant from The Robert Wood Johnson Foundation. The funding source had no role in the design, conduct, or analysis of the study; in the preparation of the manuscript; or in the decision to submit this manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PDC values for antihypertensive and antihyperlipidemic agents.

[JPG File, 41KB - jmir_v17i10e226_app1.jpg]

References

1. 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]
2. Delbanco T, Walker J, Darer JD, Elmore JG, Feldman HJ, Leveille SG, et al. Open notes: doctors and patients signing on. Ann Intern Med 2010 Jul 20;153(2):121-125. [doi: 10.7326/0003-4819-153-2-201007200-00008] [Medline: 20643992]
3. 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]
4. Ross SE, Moore LA, Earnest MA, Wittevrongel L, Lin C. Providing a Web-based online medical record with electronic communication capabilities to patients with congestive heart failure: randomized trial. J Med Internet Res 2004 May 14;6(2):e12 [FREE Full text] [doi: 10.2196/jmir.6.2.e12] [Medline: 15249261]
5. Sabaté E. Adherence to Long-Term Therapies: Evidence for Action. Geneva, Switzerland: World Health Organization; 2003. URL: http://www.who.int/chp/knowledge/publications/adherence_full_report.pdf?ua=1 [accessed 2015-06-26] [WebCite Cache ID 6ZZuAH3uM]
6. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]
7. Aitken M, Valkova S. Avoidable Costs in U.S. Healthcare: The $200 Billion Opportunity From Using Medicines More Responsibly. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2013 Jun. URL: http://www.imshealth.com/deployedfiles/imshealth/Global/Content/Corporate/IMS%20Institute/RUOM-2013/IHII_Responsible_Use_Medicines_2013.pdf [accessed 2015-06-26] [WebCite Cache ID 6ZZtRCsVu]
8. Viswanathan M, Golin CE, Jones CD, Ashok M, Blalock S, Wines RCM, et al. Volume 4. Medication Adherence Interventions: Comparative Effectiveness. Evidence Reports/Technology Assessments, No. 208. 2012 Sep. Closing the quality gap: revisiting the state of the science URL: http://effectivehealthcare.ahrq.gov/ehc/products/296/1248/EvidenceReport208_CQGMedAdherence_FinalReport_20120905.pdf [accessed 2015-06-26] [WebCite Cache ID 6ZZsb0QAc]
9. Haynes RB, Ackloo E, Sahota N, McDonald HP, Yao X. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2008(2):CD000011. [doi: 10.1002/14651858.CD000011.pub3] [Medline: 18425859]
10. Stacy JN, Schwartz SM, Ershoff D, Shreve MS. Incorporating tailored interactive patient solutions using interactive voice response technology to improve statin adherence: results of a randomized clinical trial in a managed care setting. Popul Health Manag 2009 Oct;12(5):241-254. [doi: 10.1089/pop.2008.0046] [Medline: 19848566]
11. Powell KM, Edgren B. Failure of educational videotapes to improve medication compliance in a health maintenance organization. Am J Health Syst Pharm 1995 Oct 15;52(20):2196-2199. [Medline: 8564589]
12. Schectman G, Hiatt J, Hartz A. Telephone contacts do not improve adherence to niacin or bile acid sequestrant therapy. Ann Pharmacother 1994;28(1):29-35. [doi: 10.1177/106002809402800104] [Medline: 8123955]
13. Schneider PJ, Murphy JE, Pedersen CA. Impact of medication packaging on adherence and treatment outcomes in older ambulatory patients. J Am Pharm Assoc (2003) 2008 Jan;48(1):58-63. [doi: 10.1331/JAPhA.2008.07040] [Medline: 18192132]
14. Nau DP. Pharmacy Quality Alliance. Proportion of days covered (PDC) as a preferred method of measuring medication adherence URL: http://www.pqaalliance.org/images/uploads/files/PQA%20PDC%20vs%20%20MPR.pdf [accessed 2015-06-26] [WebCite Cache ID 6ZZsCgavn]
15. Sattler EL, Lee JS, Perri M3. Medication (re)fill adherence measures derived from pharmacy claims data in older Americans: a review of the literature. Drugs Aging 2013 Jun;30(6):383-399. [doi: 10.1007/s40266-013-0074-z] [Medline: 23553512]

16. Karve S, Cleves MA, Helm M, Hudson TJ, West DS, Martin BC. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. Curr Med Res Opin 2009 Sep;25(9):2303-2310. [doi: 10.1185/03007990903126833] [Medline: 19635045]
17. Foley KA, Denke MA, Kamal-Bahl S, Simpson R, Berra K, Sajjan S, et al. The impact of physician attitudes and beliefs on treatment decisions: lipid therapy in high-risk patients. Med Care 2006 May;44(5):421-428. [doi: 10.1097/01.mlr.0000208017.18278.1a] [Medline: 16641660]
18. Martin SS, Blumenthal RS. Concepts and controversies: the 2013 American College of Cardiology/American Heart Association risk assessment and cholesterol treatment guidelines. Ann Intern Med 2014 Mar 4;160(5):356-358. [doi: 10.7326/M13-2805] [Medline: 24473832]
19. Chapman RH, Petrilla AA, Benner JS, Schwartz JS, Tang SS. Predictors of adherence to concomitant antihypertensive and lipid-lowering medications in older adults: a retrospective, cohort study. Drugs Aging 2008;25(10):885-892. [Medline: 18808213]
20. Krousel-Wood M, Joyce C, Holt E, Muntner P, Webber LS, Morisky DE, et al. Predictors of decline in medication adherence: results from the cohort study of medication adherence among older adults. Hypertension 2011 Nov;58(5):804-810 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.176859] [Medline: 21968751]
21. Sarkar U, Lyles CR, Parker MM, Allen J, Nguyen R, Moffet HH, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. Med Care 2014 Mar;52(3):194-201 [FREE Full text] [doi: 10.1097/MLR.0000000000000069] [Medline: 24374412]

Abbreviations

BL: baseline
BMI: body mass index
EHR: electronic health record
F-U: follow-up
GHS: Geisinger Health System
IQR: interquartile range
IRB: Institutional Review Board
N/A: not applicable
nC: number of included patients in the control group
nI: number of included patients in the intervention group
NNT: number needed to treat
PCP: primary care physician
PDC: proportion of days covered
RR: risk ratio

Edited by D Giordano; submitted 27.06.15; peer-reviewed by M Solomon; comments to author 21.07.15; revised version received 12.08.15; accepted 14.09.15; published 08.10.15.

Please cite as:
Wright E, Darer J, Tang X, Thompson J, Tusing L, Fossa A, Delbanco T, Ngo L, Walker J
Sharing Physician Notes Through an Electronic Portal is Associated With Improved Medication Adherence: Quasi-Experimental Study
J Med Internet Res 2015;17(10):e226
URL: http://www.jmir.org/2015/10/e226/
doi:10.2196/jmir.4872
PMID:26449757

Original Paper

Using Videogame Apps to Assess Gains in Adolescents' Substance Use Knowledge: New Opportunities for Evaluating Intervention Exposure and Content Mastery

Erika Montanaro[1*], PhD (HK), M Paed De; Lynn E Fiellin[2*], MD; Tamer Fakhouri[3*], MD; Tassos C Kyriakides[4*], PhD (HK), M Paed De; Lindsay R Duncan[5*], PhD (HK), M Paed De

[1]Yale University, New Haven, CT, United States

[2]play2PREVENT Lab at Yale, Yale School of Medicine, New Haven, CT, United States

[3]One Medical Group, San Francisco, CA, United States

[4]Yale School of Public Health, New Haven, CT, United States

[5]play2PREVENT Lab at Yale, McGill University, Montreal, QC, Canada

[*]all authors contributed equally

Corresponding Author:
Erika Montanaro, PhD (HK), M Paed De
Yale University
Yale School of Public Health, 135 College Street, Suite 200
New Haven, CT, 06510
United States
Phone: 1 203 764 4333
Fax: 1 203 764 4353
Email: erika.montanaro@yale.edu

Abstract

Background: Videogame interventions are becoming increasingly popular as a means to engage people in behavioral interventions; however, strategies for examining data from such interventions have not been developed.

Objective: The objective of this study was to describe how a technology-based intervention can yield meaningful, objective evidence of intervention exposure within a behavioral intervention. This study demonstrates the analysis of automatic log files, created by software from a videogame intervention, that catalog game play and, as proof of concept, the association of these data with changes in substance use knowledge as documented with standardized assessments.

Methods: We analyzed 3- and 6-month follow-up data from 166 participants enrolled in a randomized controlled trial evaluating a videogame intervention, PlayForward: Elm City Stories (PlayForward). PlayForward is a videogame developed as a risk reduction and prevention program targeting HIV risk behaviors (substance use and sex) in young minority adolescents. Log files were analyzed to extract the total amount of time spent playing the videogame intervention and the total number of game levels completed and beaten by each player.

Results: Completing and beating more of the game levels, and not total game play time, was related to higher substance use knowledge scores at the 3- (P=.001) and 6-month (P=.001) follow-ups.

Conclusions: Our findings highlight the potential contributions a videogame intervention can make to the study of health behavior change. Specifically, the use of objective data collected during game play can address challenges in traditional human-delivered behavioral interventions.

Trial Registration: Clinicaltrials.gov NCT01666496; https://clinicaltrials.gov/ct2/show/NCT01666496 (Archived by WebCite at http://www.webcitation.org/6cV9fxsOg)

(J Med Internet Res 2015;17(10):e245) doi:10.2196/jmir.4377

KEYWORDS

video games; intervention studies; substance use; HIV; evaluation; eHealth

Introduction

In recent decades, eHealth interventions, defined as the use of information technology in the promotion, prevention, treatment, and maintenance of health care [1], have emerged as a new and compelling form of intervention that can reach all ages and ethnic groups [2]. eHealth interventions are a promising form of intervention delivery because they (1) have the potential to reach wider audiences who are not motivated to use in-person approaches, (2) have lower delivery costs than human-delivered interventions, (3) allow for the standardization of delivery content, and (4) offer multiple dissemination channels (eg, smartphones, videogames) [2,3]. Videogames, in particular, are emerging as an effective platform for the delivery of behavioral health interventions [4-6]. For example, Kato et al [7] developed a videogame intervention to address issues related to cancer care and treatment for patients between the ages of 13 and 28 years. After playing either the intervention game or a control game for 1 hour per week for 3 months, patients in the intervention condition had significantly higher self-efficacy and knowledge related to their treatment and management of symptoms as well as greater adherence to medications [7]. This intervention used only videogame play time as an indication of exposure to the intervention. Videogames allow for an interactive experience in a virtual environment, which affords players the opportunity to experience the full spectrum of outcomes related to different choices without actually having to live through the potentially harmful consequences (eg, HIV infection) [8]. Theory-based interventions (ie, those that target and change psychological constructs) are more effective at producing behavior change [9] than non-theory-based interventions and the advancement of videogame interventions and technology in general allows researchers to better understand the entire process.

Active Ingredients of Change

Interventions designed from health behavior theory are more successful than those that are not theory-based [9,10]. Behavior change interventions are often complex, targeting multiple theoretical constructs in hopes of changing behavior [11]. The complex framework in which many interventions are built means there are a variety of factors being intervened on at once; thus, a limitation of conventional behavior change interventions is the ability to accurately catalog exposure to a given intervention component. Because behavioral interventions are now delivered via emerging technologies, opportunities to improve on traditional assessments of intervention implementation are being established [12]. Indeed, for more than a decade, researchers have been encouraged to think creatively to establish systems for gathering data through these eHealth interventions to explore the active ingredients of change [13,14]. Videogame-based interventions provide a concrete example of how innovative techniques to measure exposure can be implemented through eHealth interventions.

Importance of Implementation Guides

Videogames have the advantage of providing highly accurate and granular exposure to information through the analysis of each player's unique game play and decisions. Videogame software can create automatic logs with timestamps of all aspects of a player's experience (eg, button presses, game levels entered and exited, actions taken) to a level of detail allowing for the recreation of the entire game play session. New avenues for data collection using videogames, and broadly all Internet-delivered interventions, are emerging [11]. Implementation guides for these new assessment methods, such as using event log files, are needed to catalog these techniques in a purposeful, coherent, and understandable way to advance the field [15]. These guides will allow for the level of detail needed to create new research questions regarding game play interactions, participant experience, and process evaluation.

Substance Use

Adolescents use substances such as alcohol and marijuana. In 2011, 70% of students in grade 12 reported having tried alcohol and 50% reported being drunk at least once [16]. Marijuana is another substance used frequently in adolescence with 46% of those in grade 12 reporting using marijuana at least once in their lifetime [16]. Substance use is often associated with unprotected sexual activities [17,18] and an increased number of accidental deaths among adolescents [19]. However, it is unclear if adolescents are knowledgeable about the negative consequences related to substance use. It is critically important to examine what adolescents know about substance use to create effective risk reduction interventions.

This Study

This study acts as a guide, demonstrating (1) the analysis of automatic log files created by videogame software that catalog game play and (2) the association of these data with changes in substance use knowledge documented by assessments. This videogame-based intervention was grounded in social cognitive theory [20] and message framing [21]—derived from prospect theory [22]—and was developed to reduce HIV risk behaviors among young adolescents. The purpose of this paper is to describe how a technology-based intervention can yield meaningful, objective evidence of intervention exposure and participant experience within a behavioral intervention. As proof of concept, we will then demonstrate how these data relate to an important psychological construct related to behavioral outcomes: knowledge [23]. We sought to determine how performance in the game (eg, by points earned) relates to knowledge as measured using a standardized validated instrument. The implications of this strategy for extracting data from the game software and evaluating participants' game processes within a behavioral intervention can be extrapolated to interventions employing a wide variety of other technologies. Utilizing log files to capture a player's unique game play decisions and process will ultimately facilitate a broader understanding of behavior change and potentially support the use of videogames not only as effective interventions, but also as valid assessment tools.

Methods

Participants

A total of 333 participants, aged 11 to 14 years, were enrolled in a randomized controlled trial in which they played either an

intervention game, *PlayForward: Elm City Stories*, or a set of attention- and time-control games for a maximum of 16 hours over 6 weeks. Participants were recruited from community afterschool, school, and summer programs in New Haven, CT. All procedures were approved by the Yale University Human Investigation Committee.

The Videogame Intervention

PlayForward: Elm City Stories is a videogame developed as a risk reduction and prevention program targeting HIV risk behaviors, primarily sex- and substance use–related, in young minority adolescents. *PlayForward* was developed through extensive formative work with the target audience [5,8,24,25]. The main storyline of the game is comprised of "challenge stack" levels in which players travel through a virtual life from grade 7 to 12 and engage in role-playing scenarios where they must make decisions around risky behaviors (eg, unprotected sex, alcohol use) and experience the positive and negative consequences of those behaviors. The participants acquire risk-related knowledge, navigate peer relationships, and negotiate against peer pressure. Players encounter realistic stories experienced by middle school and high school students, such as sneaking into a significant other's house, unplanned pregnancy, vandalism, and drunk driving. Players must also earn points in mini-games designed to build knowledge or behavioral skills needed to avoid risk, such as refusal, negotiation, or peer-assessment skills. Through these mini-games, players acquire the "senses" and "powers" needed to resolve the stories (for a complete list of the mini-games see [24]). To demonstrate our analysis strategy, we will focus on *Know Power*.

Know Power provides the player with accurate and relevant information about the consequences of engaging in risky behaviors. This mini-game, comprised of 10 levels, emphasizes the development and clarification of adolescents' values and evaluations of risk behavior and provides them with information about the consequences of and alternatives to engaging in these behaviors. Positive attitudes toward avoiding substance use, delaying sexual initiation, and general risk reduction are cultivated in this component of the game. *Know Power* specifically focuses on increasing knowledge and creating positive attitudes about HIV risk reduction behaviors.

Measures

Knowledge

Participants were asked to complete a 22-item multiple-choice assessment of substance use and sexual health knowledge at each time point. The 22 items related to knowledge were adapted from several adolescent knowledge content sources, including an evidence-based curriculum that has been proven effective in reducing the risk of HIV, sexually transmitted infections, and teen pregnancy in young minority adolescents [26-28]. For this study, the subset of 8 questions addressing substance use knowledge was used. Participant response options were true, false, and not sure. These items included:

1. Drug users who use needles to inject drugs into their bodies have a greater chance of getting HIV if they share needles with other people.
2. People who use drugs occasionally can't become addicted to them.
3. Taking someone else's prescription drugs is safe as long as the person giving them to you says it is okay.
4. Using any drug can be more dangerous when taken with alcohol or other drugs.
5. People are more likely to make unsafe decisions about sex (eg, not wearing a condom) if they have been using drugs or drinking alcohol.
6. Drinking coffee or taking a cold shower can help sober someone up who is drunk.
7. If a person drinks too much alcohol, they might get sick from alcohol poisoning.
8. Alcohol poisoning can cause a person to stop breathing or choke to death on their own vomit.

Game Play Log Files

PlayForward was designed to capture granular information about each player's gameplay experience, actions, and behavior in event logs. There were 2 kinds of data collected: (1) player game state data, the traditional save/load data required so that a player can save their progress and then pick up where they left off later and (2) activity logging data that captures relevant actions players took during game play. Activity logging data included which game content players were exposed to, how long it took them to solve particular game levels, how much time was spent playing various portions of the game, when players beat particular levels of the game, and how many levels were beaten. Every action that the player performed in the game-selecting options, entering/exiting a game area, or making a choice was recorded with a timestamp. Timestamped data were stored in comma-delimited log files, which were then imported into a database for data analysis. Two game play variables from the activity logging data were extracted: the total amount of time spent playing *PlayForward* and the total number of game levels (challenge stacks) beaten. Total game play time was calculated by identifying timestamped game start and end events, calculating a time interval for each game session, and then summing the time across all play sessions during the trial. Total number of game levels beaten was calculated by adding up the number of game level completion events for each player. We did not focus on which game content players were exposed to because *PlayForward* is sequenced in a linear fashion, such that players are required to beat particular mini-game levels in a specific sequence to progress through challenge stacks. Therefore, players with the same number of total levels beaten were exposed to the same content.

Players in the *PlayForward* cohort were divided into 2 subsets (high scoring and low scoring) based on whether they were above (high scoring) or below (low scoring) the median number (median=8) of total game levels beaten. When participants were instructed to restart the game, game play time from their second attempt was factored into the total number of hours played, but completion of game levels a second time did not count toward a higher score (scores were capped at 12 because there are 12 levels in the game).

Data Analysis

Substance use knowledge was examined for participants in the *PlayForward* condition at baseline, 3 months, and 6 months postintervention. First, a repeated measures ANOVA was used to examine mastery effects on substance use knowledge over time. Bivariate relationships were also examined among log file variables and substance use knowledge. A 2-sample *t* test was used to compare substance use knowledge scores between the high- and low-scoring groups at each time point. Multivariate regression methods were used to assess the impact of baseline knowledge score, total amount of time spent playing *PlayForward*, and total number of game levels (challenge stacks) beaten by each player on the gain in substance use knowledge at 3- and 6-months postintervention.

Results

For the current study, only participants in the experimental condition (ie, *PlayForward*) were examined because these were the participants with game play data available (N=166; mean age 12.95, SD 1.03 years; mean game play time 7.27, SD 3.55 hours; median 8.24, IQR 3.70-10.14 hours; see Table 1 for demographics). Of the participants in the *PlayForward* group, 72 were in the high-scoring group, 72 were in the low-scoring group, and 22 players either had missing or corrupted log file data due to game software or transcription errors. Those 22 players with missing or corrupted log files were not included in the analyses. There was a significant difference between the age groups such that older participants were also more likely to be high-scoring players. A total of 37 players beat the *PlayForward* videogame and were instructed to restart the game and play through a second time.

Table 1. Baseline demographic characteristics of participants (N=144).

Characteristic	Group by game play score		Test statistic		*P*
	High score n=72	Low score n=72	χ^2 (*df*)	t_{142}	
Gender (female), n (%)	36 (50)	31 (43)	0.7 (1)		.40
Age (years), mean (SD)	13.18 (1.03)	12.72 (1.03)		2.68	.01
Race/Ethnicity, n (%)			0.6 (4)		.96
Caucasian	3 (4)	3 (4)			
Black	27 (38)	23 (32)			
Hispanic	25 (35)	29 (40)			
Biracial	8 (11)	8 (11)			
Other	9 (13)	9 (13)			

Group Differences in Substance Use Knowledge

Knowledge scores were assigned based on the total number of correct answers on the test (maximum score=8; mean 4.52, SD 1.81). There were no statistically significant differences between the groups with respect to baseline substance use knowledge score (high scoring: mean 4.65, SD 1.80; low scoring: mean 4.42, SD 1.82; t_{102}=–0.64, *P*=.52). Players with a high score in the game had significantly higher scores on the assessment of substance use knowledge at both 3-month (high scoring: mean 6.72, SD 1.45; low scoring: mean 4.74, SD 2.42; t_{102}=–5.11, *P*=.001) and 6-month (high scoring: mean 6.56, SD 1.53; low scoring: mean 4.78, SD 2.38; t_{102}=–4.57, *P*=.001) follow-ups. Specifically, players who mastered the intervention material (ie, high-scoring players) were also more knowledgeable of substance use facts (eg, “If a person drinks too much alcohol they might get sick from alcohol poisoning”). A significant mastery by time interaction existed for substance use knowledge ($F_{2,101}$=9.41, *P*=.001). High-scoring players had greater substance use knowledge at the 3- and 6-month follow-up periods.

Factors Associated with Substance Use Knowledge

We examined the bivariate relationships among the log file data (number of levels beaten: mean 7.64, SD 3.61; number of hours playing game: mean 7.27, SD 3.55) and substance use knowledge at baseline (mean 4.52, SD 1.81) and at the 3-month (mean 5.68, SD 2.24) and 6-month (mean 5.65, SD 2.18) follow-ups (see Table 2). Baseline substance use knowledge was not associated with either number of levels beaten or total time spent playing the game. Both 3- and 6-month substance use knowledge were positively associated with number of levels beaten or total time spent playing the game. Number of levels beaten was more strongly related to 3- and 6-month substance use knowledge than total time spent playing the game.

Table 2. Relationship between knowledge and game play variables.

Knowledge and game play	1		2		3		4		5	
	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*
1. Baseline knowledge[a]	—									
2. 3-Month knowledge[a]	.411	.001	—							
3. 6-Month knowledge[a]	.331	.001	.809	.001	—					
4. Number of levels beaten[b]	.113		.528	.001	.531	.001	—			
5. Number of hours playing game	–.033		.205	.03	.219	.02	.584	.001	—	

[a] Knowledge scores range from 0 to 8.

[b] Number of levels beaten maximum=12.

Predictors of Substance Use Knowledge at Follow-Up

Three-Month Follow-Up

Age differences were found at baseline (Table 1) such that older participants were more likely to also be high-scoring players; therefore, age was included in the model to determine if age differences were responsible for mastery of game material. Collapsing across groups (ie, high scoring vs low scoring), results of the multivariate regression analyses examining game play variables as predictors of substance use knowledge were significantly related to substance use knowledge at the 3-month follow-up (Table 3) (R^2=.396, $F_{4,102}$=18.36, *P*=.001). Only 2 predictors were significantly uniquely associated with substance use knowledge at 3 months; baseline knowledge (*B*=.281, t_{106}=3.47, *P*=.001) and game levels beaten (*B*=.489, t_{106}=5.12, *P*=.001) were positively related to substance use knowledge at 3 months. Higher scores on these constructs were associated with higher substance use knowledge at 3 months. Age and total time playing *PlayForward*, on the other hand, were not statistically significantly related to substance use knowledge at 3 months.

Table 3. Association of game play with gains in substance use knowledge.

Substance use knowledge	*B*	SE	β	t_{102}	*P*
3 Month					
Number of levels beaten	.31	0.06	.49	5.12	.001
Number of hours playing game	–.07	0.07	–.09	–1.00	.32
Baseline knowledge	.35	0.10	.28	3.47	.001
Age	.38	0.18	.16	1.85	.07
6 Month					
Number of levels beaten	.11	0.05	.18	2.19	.03
Number of hours playing game	–.02	0.05	–.03	–0.35	.71
3-month knowledge	.72	0.07	.74	10.23	.001
Age	–.13	0.13	–.06	–0.98	.33

Six-Month Follow-Up

Similar patterns appeared for substance use knowledge at the 6-month follow-up (Table 3) (R^2=.644, $F_{4,102}$=48.97, *P*=.001). Only 2 predictors were significantly uniquely associated with substance use knowledge at 6 months; 3 month knowledge (*B*=.735, t_{106}=10.23, *P*=.001) and game levels beaten (*B*=.179, t_{106}=2.19, *P*=.03) were positively related to substance use knowledge at 6 months. Higher scores on these constructs were associated with higher substance use knowledge at 6 months. Age and total time playing *PlayForward*, on the other hand, were not statistically significantly related to substance use knowledge at 6 months. Overall, it was number of levels beaten (ie, mastery of intervention content) that was associated with a gain in substance use knowledge at 3 months and 6 months.

Discussion

This study describes how a videogame intervention can produce objective data demonstrating exposure to the intervention and yield meaningful evidence of the association between a participant's success within the videogame and changes in an important psychological construct—knowledge. We examined how 2 separate measures of game play—total amount of time spent playing the *PlayForward* videogame and the total number of game levels (challenge stacks) beaten by each player—might be related to substance use knowledge at 3- and 6-month follow-ups. These 2 measures of game play broaden the scope of traditional face-to-face intervention assessment by accurately tracking the participant's success within the game. For example, it appears that mastery of the intervention material (eg, game

levels beaten), and not time spent playing each level, influences acquisition of substance use knowledge.

High-scoring game players gained and retained significantly more knowledge about substance use at the follow-up time points than their low-scoring counterparts did. Interestingly, completing and beating more of the game levels, and not total game play time, resulted in greater changes in knowledge over time. Mastery of the material increases substance use knowledge for up to 6 months after game play has been completed.

Implications for eHealth Interventions

There is a growing body of literature demonstrating eHealth interventions as efficacious tools for creating behavior change [12,29]. However, it is still largely unknown *how* this change occurs. The current study examined one behavioral antecedent of behavior change—knowledge—to illustrate how objective data collected throughout game play can contribute to our understanding of how change happens. An implicit assumption in recommendations for behavior change interventions is that more time spent in interventions will produce the greatest amount of behavior change [30]. For example, an intervention examining the effects of a single-session versus a multiple-session smoking cessation intervention suggests multiple sessions (ie, more time spent in the intervention) are related to significantly higher abstinence rates than a single session [31]. Our data, on the other hand, suggest that it is not time but mastery of the material that is related to changes of important psychological mediators (ie, knowledge) of behavior change. This is a valuable methodological point. Many eHealth interventions merely present information to participants [13], much like a brochure, but do not include content that must be mastered. Our data suggest that simple exposure to intervention content is not enough to create change; instead, what is needed is the opportunity to process intervention content more deeply. It is not enough to give people information, we must give them a means by which to master intervention content to change behavior.

Traditional human-delivered interventions rely heavily on self-report data or a measure of overall intervention time as a proxy measure to examine an individual participant's experience throughout an intervention. eHealth interventions, such as videogames, allow researchers to objectively track a participant's intervention experience through the game software by easily collecting data on time spent in individual segments of the intervention, overall time spent completing the intervention, and mastery of intervention material. This information can then be used to learn which material is most important to the behavior change process. For example, the data collection features of eHealth interventions could allow us to determine if one portion of an intervention dedicated to knowledge acquisition mastery is enough to produce knowledge gains at follow-up interval periods or if more sessions are necessary. This is just one of the questions that can be answered by using a technology-based intervention as a data collection tool. Before eHealth interventions, these questions were burdensome and resource-intensive to investigate [32]. These are empirical questions that can—and should—be answered to facilitate optimal intervention development and to inform the progress of theoretical innovations in the behavior change domain.

Behavior change interventions target multiple constructs at once to create behavior change [11]. Thus, conventional behavior change interventions are unable to determine which intervention components are contributing to behavior change. A dismantling study design could help determine if a focused manipulation of one construct can have diffuse effects on other constructs [33] and/or if such a focused manipulation of one construct is as good as or better than manipulation of multiple constructs simultaneously. Given the lower delivery cost, standardization of delivery content, multiple dissemination channels, and objective data collection of participant experience [2,3], eHealth interventions are in a unique and prime position to answer these questions.

Limitations

With this study, as a demonstration of the novel methods of evaluating gameplay data, we chose to examine the behavioral antecedent of knowledge, not behavioral outcomes, which will be included in future analyses. Another limitation of this study was that knowledge relied on a self-report measure; this limitation is shared with the majority of intervention research. However, log file data offer unique assessment techniques to examine intervention exposure and outcomes whereby adding another valid source of data and potentially reducing the need to rely solely on self-report data. Additionally, only a small number of items (ie, 8) were used to assess substance use knowledge.

Conclusions

Increasingly, eHealth interventions, particularly videogames, are becoming a standard and effective form of intervention delivery [2-5]. Interventions delivered via emerging technologies have the potential to enhance traditional forms of assessment with real-time objective data collection techniques. Importantly, videogames can and do serve a dual purpose of delivering health interventions and providing a unique assessment tool with minimal additional effort required from participants. Participant log files are just one potential technique. Our findings highlight the valuable contributions eHealth interventions make to the study of health behavior change by using objective data collection techniques that are difficult to replicate in traditional human-delivered behavioral interventions. This research illustrates the need for the rigorous exploration of new data collection opportunities provided by eHealth interventions. These innovative methods hold the potential to document specific engagement, interactions, and exposure that may provide key information regarding which intervention components are most effective in influencing participant outcomes.

Acknowledgments

This research was supported by NICHD R01 HD062080-01 (ClinicalTrials.gov: NCT01666496) from the National Institute of Child Health and Human Development (PI: Lynn E Fiellin, MD). Dr Montanaro is supported by T32MH020031 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health. Additional thanks to Kimberly Hieftje, PhD and Tyra Pendergrass, MEM for their assistance with data collection.

Conflicts of Interest

Drs Fiellin and Duncan are affiliated with KnackTime Interactive, a small commercial venture that focuses on the distribution of evidence-based videogames for risk reduction and prevention in youth and young adults. This relationship is extensively managed by Drs Fiellin and Duncan and their respective institutions.

References

1. 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]
2. Lehto T, Oinas-Kukkonen H. Persuasive features in web-based alcohol and smoking interventions: A systematic review of the literature. J Med Internet Res 2011;13(3):e46 [FREE Full text] [doi: 10.2196/jmir.1559] [Medline: 21795238]
3. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: A meta-analysis. AIDS 2009 Jan 2;23(1):107-115. [doi: 10.1097/QAD.0b013e32831c5500] [Medline: 19050392]
4. Hieftje K, Edelman EJ, Camenga DR, Fiellin LE. Electronic media-based health interventions promoting behavior change in youth: A systematic review. JAMA Pediatr 2013 Jun;167(6):574-580 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.1095] [Medline: 23568703]
5. Fiellin LE, Hieftje KD, Duncan LR. Videogames, here for good. Pediatrics 2014 Nov;134(5):849-851. [doi: 10.1542/peds.2014-0941] [Medline: 25287452]
6. Ruggiero L, Moadsiri A, Quinn LT, Riley BB, Danielson KK, Monahan C, et al. Diabetes island: Preliminary impact of a virtual world self-care educational intervention for African Americans with type 2 diabetes. JMIR Serious Games 2014;2(2):e10 [FREE Full text] [doi: 10.2196/games.3260] [Medline: 25584346]
7. 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]
8. Hieftje K, Duncan LR, Fiellin LE. Novel methods to collect meaningful data from adolescents for the development of health interventions. Health Promot Pract 2014 Sep;15(5):714-722 [FREE Full text] [doi: 10.1177/1524839914521211] [Medline: 24519998]
9. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
10. Noar SM. Behavioral interventions to reduce HIV-related sexual risk behavior: Review and synthesis of meta-analytic evidence. AIDS Behav 2008 May;12(3):335-353. [doi: 10.1007/s10461-007-9313-9] [Medline: 17896176]
11. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. Am J Prev Med 2007 Oct;33(4):336-345 [FREE Full text] [doi: 10.1016/j.amepre.2007.05.007] [Medline: 17888860]
12. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: Are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
13. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]
14. Neuhauser L, Kreps GL. Rethinking communication in the e-health era. J Health Psychol 2003 Jan;8(1):7-23. [doi: 10.1177/1359105303008001426] [Medline: 22113897]
15. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
16. Johnston L, O'Malley P, Bachman J, Schulenberg J. Monitoring the Future National Survey Results on Drug Use, 1975–2011. Volume I: Secondary School Students. Ann Arbor, MI: University of Michigan; 2012.
17. Weinhardt LS, Carey MP. Does alcohol lead to sexual risk behavior? Findings from event-level research. Annu Rev Sex Res 2000;11:125-157 [FREE Full text] [Medline: 11351830]
18. Anderson BJ, Stein MD. A behavioral decision model testing the association of marijuana use and sexual risk in young adult women. AIDS Behav 2011 May;15(4):875-884 [FREE Full text] [doi: 10.1007/s10461-010-9694-z] [Medline: 20358274]
19. Eaton D, Kann L, Kinchen S, Shanklin S, Flint KH, Hawkins J, Centers for Disease Control and Prevention. Youth risk behavior surveillance—United States, 2011. MMWR Morb Mortal Wkly Rep 2012;61(4):1-162.

20. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998;13(4):623-649.
21. Rothman AJ, Salovey P. Shaping perceptions to motivate healthy behavior: The role of message framing. Psychol Bull 1997 Jan;121(1):3-19. [Medline: 9000890]
22. Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. Econometrica 1979;263:291.
23. Fisher JD, Fisher WA, Misovich SJ, Kimble DL, Malloy TE. Changing AIDS risk behavior: Effects of an intervention emphasizing AIDS risk reduction information, motivation, and behavioral skills in a college student population. Health Psychol 1996 Mar;15(2):114-123. [Medline: 8681919]
24. Duncan LR, Hieftje KD, Culyba S, Fiellin LE. Game playbooks: Tools to guide multidisciplinary teams in developing videogame-based behavior change interventions. Transl Behav Med 2014 Mar;4(1):108-116 [FREE Full text] [doi: 10.1007/s13142-013-0246-8] [Medline: 24653781]
25. Hieftje K, Rosenthal MS, Camenga DR, Edelman EJ, Fiellin LE. A qualitative study to inform the development of a video game for adolescent HIV prevention. Games Health J 2012 Aug 10;1(4):294-298 [FREE Full text] [doi: 10.1089/g4h.2012.0025] [Medline: 24078897]
26. Jemmott JB, Jemmott LS, Fong GT. Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. Arch Pediatr Adolesc Med 2010 Feb;164(2):152-159 [FREE Full text] [doi: 10.1001/archpediatrics.2009.267] [Medline: 20124144]
27. Kelly JA, St Lawrence JS, Hood HV, Brasfield TL. An objective test of AIDS risk behavior knowledge: scale development, validation, and norms. J Behav Ther Exp Psychiatry 1989 Sep;20(3):227-234. [Medline: 2632590]
28. Jemmott L, Jemmott J, McCaffree K. Making a Difference!: An Abstinence Approach to HIV/STD and Teen Pregnancy Prevention. New York: Select Media, Inc; 2002:1-263.
29. 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]
30. Peters E, Klein W, Kaufman A, Meilleur L, Dixon A. More is not always better: intuitions about effective public policy can lead to unintended consequences. Soc Issues Policy Rev 2013 Jan 1;7(1):114-148 [FREE Full text] [doi: 10.1111/j.1751-2409.2012.01045.x] [Medline: 24000291]
31. Zhu SH, Stretch V, Balabanis M, Rosbrook B, Sadler G, Pierce JP. Telephone counseling for smoking cessation: Effects of single-session and multiple-session interventions. J Consult Clin Psychol 1996 Feb;64(1):202-211. [Medline: 8907100]
32. Jacobson NS, Dobson KS, Truax PA, Addis ME, Koerner K, Gollan JK, et al. A component analysis of cognitive-behavioral treatment for depression. J Consult Clin Psychol 1996 Apr;64(2):295-304. [Medline: 8871414]
33. Borkovec TD, Newman MG, Pincus AL, Lytle R. A component analysis of cognitive-behavioral therapy for generalized anxiety disorder and the role of interpersonal problems. J Consult Clin Psychol 2002 Apr;70(2):288-298. [Medline: 11952187]

Edited by G Eysenbach; submitted 24.02.15; peer-reviewed by L Ruggiero, W Nilsen; comments to author 23.07.15; revised version received 03.08.15; accepted 15.08.15; published 28.10.15.

Please cite as:
Montanaro E, Fiellin LE, Fakhouri T, Kyriakides TC, Duncan LR
Using Videogame Apps to Assess Gains in Adolescents' Substance Use Knowledge: New Opportunities for Evaluating Intervention Exposure and Content Mastery
J Med Internet Res 2015;17(10):e245
URL: http://www.jmir.org/2015/10/e245/
doi:10.2196/jmir.4377
PMID:26510775

Original Paper

Are Health-Related Tweets Evidence Based? Review and Analysis of Health-Related Tweets on Twitter

Khalid A Alnemer[1*], MD; Waleed M Alhuzaim[2*], MD; Ahmed A Alnemer[2*], MD; Bader B Alharbi[3*], MBBS; Abdulrahman S Bawazir[3*], MBBS; Omar R Barayyan[3*], MBBS; Faisal K Balaraj[3*], MBBS

[1]Al-Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

[2]Department of Internal Medicine, Al-Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

[3]Kind Saud University, Riyadh, Saudi Arabia

[*]all authors contributed equally

Corresponding Author:
Khalid A Alnemer, MD
Department of Internal Medicine
Al-Imam Mohammad Ibn Saud Islamic University
Riyadh, 13317-4233.
Saudi Arabia
Phone: 966 1 2037100
Fax: 966 1 2590209
Email: alnemerk@hotmail.com

Abstract

Background: Health care professionals are utilizing Twitter to communicate, develop disease surveillance systems, and mine health-related information. The immediate users of this health information is the general public, including patients. This necessitates the validation of health-related tweets by health care professionals to ensure they are evidence based and to avoid the use of noncredible information as a basis for critical decisions.

Objective: The aim of this study was to evaluate health-related tweets on Twitter for validity (evidence based) and to create awareness in the community regarding the importance of evidence-based health-related tweets.

Methods: All tweets containing health-related information in the Arabic language posted April 1-5, 2015, were mined from Twitter. The tweets were classified based on popularity, activity, interaction, and frequency to obtain 25 Twitter accounts (8 physician accounts, 10 nonofficial health institute accounts, 4 dietitian accounts, and 3 government institute accounts) and 625 tweets. These tweets were evaluated by 3 American Board–certified medical consultants and a score was generated (true/false) and interobserver agreement was calculated.

Results: A total of 625 health-related Arabic-language tweets were identified from 8 physician accounts, 10 nonofficial health institute accounts, 4 dietician accounts, and 3 government institute accounts. The reviewers labeled 320 (51.2%) tweets as false and 305 (48.8%) tweets as true. Comparative analysis of tweets by account type showed 60 of 75 (80%) tweets by government institutes, 124 of 201 (61.7%) tweets by physicians, and 42 of 101 (41.6%) tweets by dieticians were true. The interobserver agreement was moderate (range 0.78-0.22). More than half of the health-related tweets (169/248, 68.1%) from nonofficial health institutes and dietician accounts (59/101, 58.4%) were false. Tweets by the physicians were more likely to be rated “true” compared to other groups (P<.001).

Conclusions: Approximately half of the medical tweets from professional accounts on Twitter were found to be false based on expert review. Furthermore, most of the evidence-based health-related tweets are posted by government institutes and physicians.

(J Med Internet Res 2015;17(10):e246) doi:10.2196/jmir.4898

KEYWORDS

health; diseases; daily medical information; medical accounts; health accounts; doctor accounts; nutrition accounts

Introduction

Twitter it is a free social networking website established in July 2006, which enables users to write and read online posts (known as "tweets") that are limited to 140 characters. Tweets can be posted via the Web, instant online message, or mobile phone. Twitter has more than 500 million active users who generate more than 340 million tweets and 1.6 billion search queries per day.

Several health care professionals in the Middle Eastern countries use social media, especially Twitter, because of its ability to connect seamlessly with colleagues, patients, and other medical professionals. It is also a great resource to educate the public, track disease outbreaks, collect real-time health data, recruit study participants, recognize misuse of antibiotics, and gain more knowledge on health-related topics.

Several studies have examined the content of health-related tweets on Twitter. A study that investigated all posts with the words "Ebola" and "prevention" or "cure" from Guinea, Liberia, and Nigeria showed that the most common misinformation was that Ebola might be cured by the plant ewedu or by blood transfusion [1]. A study in 2010 investigated status updated from 52,153 tweets with the combination "flu + antibiotics" and "cold + antibiotics" associated with misinformation. Results showed a total of 172,571 and 850,375 followers of misinformation, respectively, for the 2 combinations [2].

Tweets were primarily used to disseminate information from credible sources, but were also a source of opinions and experiences [3]. A study conducted in Norway regarding the content and seriousness of tweets on chlamydia and HIV showed that 9 of 10 tweets on HIV were of serious nature and many of the tweets that were retweeted were facts [4]. A study conducted to evaluate opinions and knowledge regarding computed tomography radiation risk from 621 tweets posted by 557 accounts (doctors: 16%; health institute: 5%; patients: 6%; technologists: 1%; other users: 71%) showed that most tweets were not peer-reviewed, were posted by nonphysicians, and content was unfavorable [5].

A study conducted in 2014 at The John Hopkins University to analyze the content of 665 tweets on Twitter showed that 346 were health-related tweets, 53.2% were testable claims, 41.0% were news, 26.9% were commercial product or service, 17.6% were personal experience, and 17.1% were about wellness [6].

These studies indicate that the validity of health-related tweets on Twitter needs to be assessed, especially to check if the content represents a claim supported by evidence, a personal opinion, or other information. Thus, this study reports the results of a content analysis of health-related tweets on Twitter in Arabic.

Methods

Twitter was chosen to investigate health-related tweets because it is the most common social media in the Gulf countries. Only tweets in Arabic were included in this study.

A manual approach was used to identify and categorize health-related tweets posted by health-related accounts, associated with either an organization or an individual user. These tasks were crucial for the identification, data collection, and categorization process for this study.

Identification of Relevant Twitter Accounts

The relevant accounts were identified via a 4-step process. The first step involved a search of the Twitter website using the following search terms in Arabic: health, your health, agility, regimen, healthy diet, drugs , disease, diseases, drug, treatment, prohibited drugs, epidemic, inflammations, infection, medical information, doctors, hospitals, daily medical information, nutrition, medical accounts, health accounts, doctor accounts, and nutrition accounts (see Multimedia Appendix 1 for Arabic search terms). This search resulted in 203 tweets that were reviewed; those accounts whose identity could not be ascertained were excluded.

The second step involved selection of accounts based on:

1. Number of followers (minimum number was set to 250,000);
2. Activity (tweeted for the period of April 2015);
3. Interaction with other users; and
4. Frequency of tweets (health-related tweets on a daily basis).

This resulted in a list of 86 Twitter accounts: 31 physician accounts, 39 nonofficial health institute accounts, 6 dietitian accounts, 2 media accounts, and 8 government institute accounts.

The third step involved further examination of the 86 accounts by the following 3 criteria:

1. Popularity (most viewed);
2. Interaction with other users; and
3. Number of followers (minimum number was set to 45,000).

Accounts with a minimum of 45,000 followers were reviewed for 1 week to select those posting at least 5 health-related tweets per day with at least 100 retweets a week. This resulted in a list of 25 Twitter accounts: 8 physician accounts, 10 nonofficial health institute accounts, 4 dietitian accounts, and 3 government institute accounts.

Physician Twitter Accounts

Twitter accounts were identified as those whose description provided a Web link to their corresponding clinics/hospital website. In total, there were 8 physician Twitter accounts.

Nonofficial Health Institute Twitter Accounts

Twitter accounts that were identified as those whose description provided a Web link to a nonofficial health institute. These accounts had a range of 70,000 to 300,000 followers. There were a total of 10 such accounts.

Dietician Twitter Accounts

Twitter accounts that were identified as those whose description provided a Web link to their clinic/hospital website. These accounts had a range of 45,000 to 210,000 followers. There were a total of 4 such accounts.

Government Institute Twitter Accounts

Twitter accounts that were identified as those whose description provided a Web link to their corresponding governmental site (ie, ending with gov.sa). These accounts had a range of 43,000 to 843,000 followers. There were a total of 3 such accounts.

The final step involved selection of the first 5 health-related tweets daily for 5 days (April 1-5, 2015) from each of these 25 accounts. This resulted in a total of 625 tweets, which were integrated in a Microsoft Word file.

Examination of Individual Tweets

The Word file was evaluated by 3 independent reviewers (American Board–certified consultants with more than 10 years of experience in medical practice collaborating with other specialty consultants in different fields, if needed) who were blind to the identity of the Twitter users during content analysis. The reviewers evaluated and labeled these tweets as false, true with weak evidence (ie, expert opinion), true with moderate evidence (ie, small randomized controlled trial [RCT], nonrandomized observational study, registry), or true with strong evidence (ie, many large RCTs).

This was followed by scoring of the tweets, a system that used the majority of the reviewer's opinions to generate a score for each tweet. For example, if 2 of 3 experts decided on moderate evidence, moderate evidence was chosen as the score for the tweet.

If there was no majority in the reviewers' opinion, the lower evidence level was chosen as the score. For example, if the 3 reviewers chose weak evidence, moderate evidence, and false, respectively, because the majority had ranked it as true, weak evidence was chosen as the score for the tweet.

Descriptive statistics were used to tabulate types of account and response of each reviewer. Comparative analysis of type of Twitter account and chi-square tests were used to determine statistical significance of the result. Interobserver agreement for the 3 independent reviewers was based on the following formula: (true/[true+false]).

Results

The data collection process for this study is presented in Figure 1. Overall, 625 Arabic-language health-related tweets contributed by 25 user accounts were analyzed as defined in Table 1.

Table 1. Tweets by type of account (total tweets: N=625).

Account	n (%)
Physician	201 (32.2)
Government institute	75 (12.0)
Nonofficial health institute	248 (39.7)
Dietician	101 (16.2)

The evaluation of each health-related tweet and categorization into 1 of 4 categories (false, weak, moderate, or strong) by the 3 independent reviewers. In the absence of a majority within the true category, weak evidence was chosen as the score for the tweet (Table 2).

Table 2. Coding of tweets by reviewer.

Reviewers decision	False, n (%)	True, n (%)			Interobserver agreement
		Weak	Moderate	Strong	
Expert 1	268 (42.9)	332 (53.1)	20 (3.2)	5 (0.8)	0.57
Expert 2	140 (22.4)	226 (36.2)	173 (27.7)	86 (13.8)	0.78
Expert 3	488 (78.1)	62 (9.9)	72 (11.5)	3 (0.5)	0.22
Final decision	320 (51.2)	261 (41.8)	39 (6.2)	5 (0.8)	

More than half of the tweets (320/625, 51.2%) in this sample were not supported by medical evidence (Table 2). The interobserver agreement between the 3 independent reviewers ranged from 0.78 to 0.22 (Table 2).

Comparative analysis of tweets by account type showed that 60 of 75 (80%) tweets by government institutes, 124 of 201 (61.7%) tweets by physicians, and 42 of 101 (41.6%) tweets by dieticians were true.

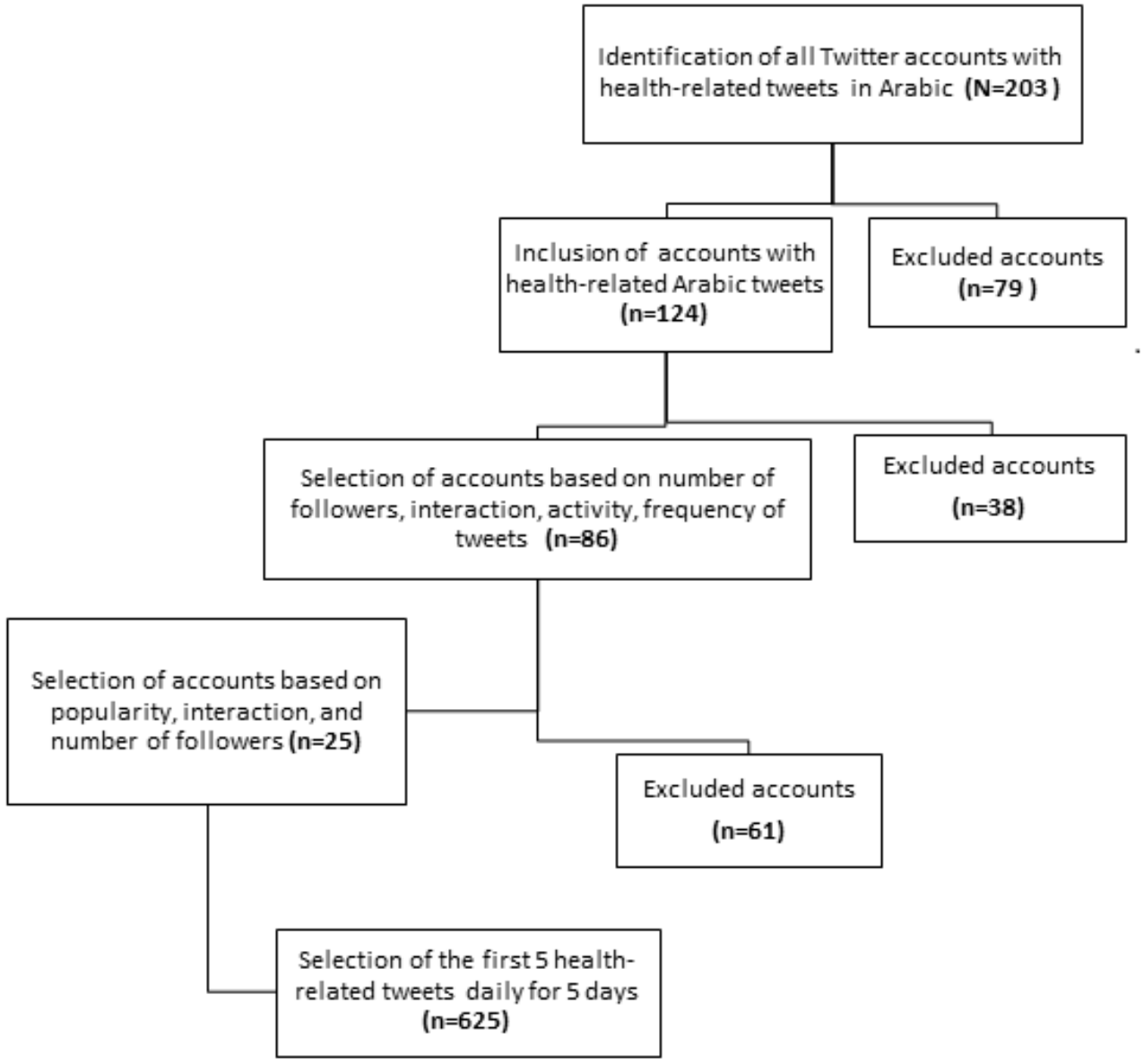
More than half of the health-related tweets from nonofficial health institutes (169/248, 68.1%) and dietician accounts (59/101, 58.4%) were false. Tweets by the physicians were more likely to be rated as "true" compared to other groups (P<.001) (Table 3).

Table 3. Comparative analysis of account type and final validity of tweets.

Type of account	Final opinion, n (%)[a]	
	False n=320	True n=305
Government institute	15 (20.0)	60 (80.0)
Physician	77 (38.3)	124 (61.7)
Nonofficial health institute	169 (68.1)	79 (31.9)
Dietician	59 (58.4)	42 (41.6)

[a] For 4×2 table, $P<.001$.

Figure 1. Collection of data flowchart.



Discussion

A comparative analysis of the 625 health-related Arabic-language tweets showed that 60 of 75 (80%) tweets by government institutes, 124 of 201 (61.7%) tweets by physicians, and 42 of 101 (41.6%) tweets by dieticians were evidence based. More than half of the health-related tweets (169/248, 68.1%) from nonofficial health institutes were false and tweets by the physicians were more likely to be rated as true compared to other groups ($P<.001$).

Twitter is an online minefield of health-related information, which can considerably affect patient health. It allows for

seamless patient-physician relationships and access to infinite online discussions and information on health-related topics.

With the advent of social media, health advice and recommendations from the Internet can influence a patient apart from their physician. Patients are becoming increasingly aware of treatment options, health literacy, and knowledge about disease. Health-related information online is especially beneficial for patients who are immobile and homebound as a result of debilitating illness [7]. However, the Web can also be used to foster unscientific health messages; therefore, patients who are unable to distinguish valid tweets from invalid ones may be misguided. Thus, health care professionals need to be vigilant and responsible for health-related information posted on the Internet [8].

Furthermore, a group needs to be created including government institutes, physicians, other health care professionals, and researchers to ensure that online health care resources are current, credible, and reliable for patient use. This information should be available in a format that is user-friendly, comprehensible, and easily accessible. The use of valid evidence-based Web resources can ensure patient-friendly formats [8,9].

This is the first study to review health-related tweets posted in Arabic language and is comparable to other reviews of health-related reviews in English [10]. The results of this study show that the content of tweets by health-related users on Twitter varied with user type (ie, the government institutes share most of evidence-based medical tweets). Also, in this study 1 in 3 physicians shared health information that was rated false in contrast with previous studies that showed physicians shared testable claims.

This study also shows the reviewers achieved moderate levels of agreement (0.78-0.22) in the classification of tweets as true or false. This could be attributed to expert 2, whose “true” votes were higher than the other 2 experts. This highlights the need for further discussions between the experts regarding classifying the health-related tweets.

Limitations

The selection of the sample size was limited by health-related tweets in Arabic language on Twitter; thus, it was not feasible to select a random sample of all health users and their tweets reducing the generalizability of the study results.

This study included categorization of health-related tweets by user account followed by analysis. However, the user category designated as “nonofficial health institutes” consisted of accounts whose background was not verifiable and may not be from health-related users. Thus, the results likely underestimate potential differences between groups, emphasizing the need for in-depth analysis.

In addition, user accounts were neither verified independently nor checked against other databases. Furthermore, the low interobserver agreement is attributed to each reviewer’s perception of evidence based, rather than based on evidence, which limits the validity of the results.

This study gives two clear and simple messages to health care professionals, patients, and the general public who access Arabic medical tweets. Firstly, the medical information obtained from accounts on Twitter needs to be confirmed with evidence before applying to real-life situations. Secondly, the scientific value of the tweets from government institutes and physicians is higher compared with other users.

Future Recommendations

The findings of this study set a baseline for future analyses. Our study recommends developing a consensus around the types of tweets physicians should send based on a minimal evidence level that should be included in the tweet. In addition, health care professionals need to work toward creation of guidelines and policies on the use of social media in modern health care.

Conclusions

Approximately half of the medical tweets from professional accounts on Twitter were found to be false based on expert review. Furthermore, most of the evidence-based health-related tweets were posted by government institutes and physicians.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Arabic search terms.

[PDF File (Adobe PDF File), 180KB - jmir_v17i10e246_app1.pdf]

References

1. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? BMJ 2014;349:g6178. [Medline: 25315514]
2. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. Am J Infect Control 2010 Apr;38(3):182-188 [FREE Full text] [doi: 10.1016/j.ajic.2009.11.004] [Medline: 20347636]
3. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]

4. Gabarron E, Serrano JA, Wynn R, Lau AY. Tweet content related to sexually transmitted diseases: no joking matter. J Med Internet Res 2014;16(10):e228 [FREE Full text] [doi: 10.2196/jmir.3259] [Medline: 25289463]
5. Prabhu V, Rosenkrantz A. AuntMinnie.com. 2015. Twitter posts highlight patient fears over CT radiation risk URL: http://www.auntminnie.com/index.aspx?sec=ser&sub=def&pag=dis&ItemID=109803
6. Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD. What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. J Med Internet Res 2014;16(10):e237 [FREE Full text] [doi: 10.2196/jmir.3765] [Medline: 25591063]
7. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
8. Tonsaker T, Bartlett G, Trpkov C. Health information on the Internet: gold mine or minefield? Can Fam Physician 2014 May;60(5):407-408 [FREE Full text] [Medline: 24828994]
9. Eysenbach G, Diepgen TL. The role of e-health and consumer health informatics for evidence-based patient choice in the 21st century. Clin Dermatol 2001;19(1):11-17. [Medline: 11369478]
10. Eysenbach G. Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. J Med Internet Res 2011;13(4):e123 [FREE Full text] [doi: 10.2196/jmir.2012] [Medline: 22173204]

Abbreviations

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 03.07.15; peer-reviewed by M Larsen, M Decamp, B Liang, S Mooijaart; comments to author 03.08.15; revised version received 06.09.15; accepted 07.10.15; published 29.10.15.

Please cite as:
Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, Balaraj FK
Are Health-Related Tweets Evidence Based? Review and Analysis of Health-Related Tweets on Twitter
J Med Internet Res 2015;17(10):e246
URL: http://www.jmir.org/2015/10/e246/
doi:10.2196/jmir.4898
PMID:26515535

Original Paper

A Computerized Lifestyle Application to Promote Multiple Health Behaviors at the Workplace: Testing Its Behavioral and Psychological Effects

Sonia Lippke[1,2], Dipl Psych, PhD (Psych); Lena Fleig[3,4,5], Dipl Psych, PhD (Psych); Amelie U Wiedemann[3], Dipl Psych, PhD (Psych); Ralf Schwarzer[3,6,7], Dipl Psych, PhD (Psych)

[1]Jacobs Center for Lifelong Learning and Institutional Development (JCLL), Focus Area Diversity, Health Psychology, Jacobs University Bremen, Bremen, Germany

[2]Bremen International Graduate School for Social Sciences (BIGSSS), Bremen, Germany

[3]Department of Education and Psychology, Freie Universitaet Berlin, Berlin, Germany

[4]Centre for Hip Health and Mobility, Vancouver, BC, Canada

[5]Department of Family Practice, University of British Columbia, Vancouver, BC, Canada

[6]Institute for Positive Psychology and Education, Australian Catholic University, Strathfield, Australia

[7]University of Social Sciences and Humanities, Wroclaw, Poland

Corresponding Author:
Sonia Lippke, Dipl Psych, PhD (Psych)
Jacobs Center for Lifelong Learning and Institutional Development (JCLL)
Focus Area Diversity, Health Psychology
Jacobs University Bremen
Campus Ring 1
Bremen, 28759
Germany
Phone: 49 421 200 ext 4730
Fax: 49 421 200 49 4730
Email: s.lippke@jacobs-university.de

Abstract

Background: Preventive health behaviors, such as regular physical activity and healthy nutrition, are recommended to maintain employability and to facilitate the health of employees. Theory-based workplace health promotion needs to include psychological constructs and consider the motivational readiness (so-called stages of change) of employees. According to the stages, people can be grouped as nonintenders (not motivated to change and not performing the goal behavior), intenders (decided to adopt the goal behavior but not started yet), or actors (performing the goal behavior already). The tailoring to these stages can be done computer based and should make workplace health promotion more effective.

Objective: It was tested whether a parsimonious computer-based health promotion program implemented at the workplace was effective in terms of lifestyle changes and psychological outcomes as well as body weight. We hypothesized that the stage-matched intervention would outperform the one-size-fits-all active control condition (standard care intervention).

Methods: In a randomized controlled trial, a total of 1269 employees were recruited by a trained research assistant at their workplace during a routine medical examination. After excluding noneligible employees, 560 completed Time 1 (T1), and 384 also completed Time 2 (T2), achieving a retention rate of 68.6%. Two fully automated computer-based treatments were adopted: (1) an active control condition with information about benefits of exercise and healthy nutrition (n=52), or (2) a stage-matched multiple-behavior intervention that provided different psychological treatments to 9 subgroups, addressing stages of change (nonintenders, intenders, and actors per behavior; n=332). Baseline assessments (T1) on behavior, psychological constructs, and body weight were repeated after 4 weeks (T2).

Results: The stage-matched intervention outperformed the active control condition for lifestyle changes containing physical activity and nutrition (χ^2_1=3.5; *P*=.04, for N=384) as well as psychological variables (physical activity intention, *P*=.04; nutrition intention, *P*=.03; nutrition planning, *P*=.02; and general social support to live healthily, *P*=.01). When predicting a healthy lifestyle at follow-up, baseline lifestyle (odds ratio, OR, 2.25, 95% CI 1.73-2.92; *P*<.01) and the intervention (OR 1.96, 95% CI 1.00-3.82;

P=.05) were found to be significant predictors. Physical activity planning mediated the effect of the intervention on the adoption of an overall healthy lifestyle (consisting of activity and nutrition, R^2_{adj}=.08; *P*<.01), indicating that if the stage-matched intervention increased planning, the adoption of a healthy lifestyle was more likely.

Conclusions: Matching an intervention to the motivational readiness of employees can make a health promotion program effective. Employees' motivation, planning, social support, and lifestyle can be supported by a stage-matched intervention that focuses on both physical activity and healthy nutrition. Occupational settings provide a potential to implement parsimonious computer-based health promotion programs and to facilitate multiple behavior change.

(J Med Internet Res 2015;17(10):e225) doi:10.2196/jmir.4486

KEYWORDS

demands; nutrition; physical activity; planning; workplace health promotion

Introduction

Background

Many employees wonder how they can stay healthy and maintain their employability. Further to this, many employers and organizations have to deal with preventing absenteeism (being absent due to illness) and presenteeism (being at work but not working efficiently) [1-3]. Employability and absenteeism are related to healthy lifestyles of employees, and the healthy lifestyle not only prevents physical health issues but also aids in coping with stressors at work [4]. Costs related to loss in productivity and absenteeism have been found to be associated with excess body weight: employees with a higher body mass index (BMI) were more likely to exhibit more annual sick leave days [5]. Specifically, obese employees, in comparison to employees with normal weight, had 3 or more excess sick leave days. According to a recent study, the extrapolated excess costs for employees with obesity in Germany amount to €2.18 billion [5]. These costs for Canada, also attributed to employee obesity, were estimated to be US $4.3 billion [6]. Lehnert et al [5] concluded that this calls for improved health promotion efforts. This study aims at testing an individual-level health promotion program that targets human factors, namely, health behaviors and their psychological antecedents.

Workplace Health Promotion Programs, Obesity, and Health Behaviors

In various studies, substantial proportions of the total cost of productivity loss due to sick leave and disability pensions were attributed to obesity and obesity-related diseases [7]. Thus, it is imperative to find ways to improve the health status and to lower obesity rates in the workforce. Body weight reduction can be addressed not only by physical activity but also by dietary changes [8]. Workplace health promotion programs addressing different health behaviors are promising: employees performing regular physical activity and eating healthy are less likely to exhibit a loss in productivity, even if the BMI does not decrease [9].

A Cochrane Systematic Review [10] evaluated interventions that have addressed both behaviors and examined them repeatedly over up to 24 months. Although no clear evidence for improvements in BMI could be found, physical activity as well as fruit and vegetable consumption increased. A meta-analysis of 18 studies on the efficacy of workplace health promotion [11] addressing different health behaviors found that the overall effect on work productivity and work ability was small but significant (effect sizes 0.41-0.54; *P*=.05).

Mastellos et al [10] concluded that very few studies addressing both behaviors at the same time existed, that their methodological quality was limited, and that outcomes were reported inadequately. With regard to workplace health promotion programs, Rongen et al [11] arrived at a similar conclusion. Thus, further studies with higher methodological quality and different outcomes should be conducted. Besides testing BMI and behavior change as outcomes, predictors of health behavior change should also be scrutinized; predictors such as intention, planning, and social support have been found to impact behavior change [12]. The aim of this study, therefore, was to employ a randomized control design and to use an active control group for comparison (standard care intervention) with the intervention group instead of a no-treatment group. Specifically, this design examined a lifestyle intervention addressing 2 health behaviors, namely, nutrition and physical activity.

A recent review on health promotion interventions implemented by occupational health services that aimed at physical activity and/or dietary behavior found promising effects [13]. The authors of that review concluded that counseling interventions targeting at-risk individuals were successful. However, counseling by face-to-face interventions is resource demanding. In addition, because face-to-face interventions are difficult to conduct if employees cannot attend such a counseling appointment in person because of various reasons (eg, night shift or remote workers), applying computer-based counseling appears to be a good alternative.

Computer-Based Interventions and Matching of Treatments

Computer-based technology bears the advantage of having a better reach and allows greater flexibility for employees. Computer-based interventions that target health behaviors have been designed and tested over the last decades. Computer- and Internet-based interventions offer options for tailoring interventions to the needs of the individuals. A substantial body of research has shown the efficacy of tailored programs administered via print, Internet, local computer/kiosk, and telephone. An impact was not only proven on dietary change and physical activity, but also on multiple behavioral changes

[14]. A Cochrane Systematic Review [15] found that computer-based interventions led to more weight loss and limited weight regain as compared with minimal interventions.

To design interventions successfully, a useful method was to match interventions to the individuals' needs, a strategy known as "stage matching." Individual-level workplace health interventions may be matched to participants' individual stages or readiness to change [14] based on stage theories. Stage theories propose that individuals pass through different stages on their way toward behavior change. At different stages, people exhibit different mind-sets delineated by differences in their intention, action plans, coping plans, and levels of behavioral performance [16]. This implies that interventions can be matched to a person's stage of behavioral change by targeting stage-specific needs, as opposed to "one-size-fits-all" treatments or generic communication [17].

Theoretical Backdrop of Stage Matching

The health action process approach (HAPA) [16], which served as the theoretical backdrop for the development of the current workplace health intervention, distinguishes between the following 3 stages: (1) a "nonintention stage," including persons (nonintenders) who have not (yet) set the goal to act according to a previously defined criterion; (2) an "intention stage," comprising individuals motivated to change, but not yet acting (intenders); and (3) an "action stage," including persons who have already attained the behavioral criterion (actors). With that, the HAPA is parsimonious as it considers previous behavior performance and motivation to change in the future.

The HAPA proposes that nonintenders must first increase their motivation and set the goal toward changing their behavior. Risk awareness and outcome expectancies are crucial in this process. As soon as people have set the goal, they become intenders and must plan how to initiate a behavioral change. In general, social support is crucial for maintaining successful behavior change. While social support should be addressed mainly in actors, it should also be increased in all individuals who actually adopt the new behavior. According to the HAPA, coping plans support intenders as well as actors in maintaining their (recently initiated) activity levels (coping planning includes anticipation of barriers and planning what to do when facing those barriers to ensure goal pursuit). There is some experimental evidence that attests the differential efficacy of HAPA stage-matched interventions in persons with different baseline characteristics [18-21]. However, no evidence regarding employees can be found, and therefore, this study is supposed to fill this gap.

Aim and Hypotheses

The main research aim of this study was to test the efficacy of a stage-matched intervention in comparison with an active control condition (one-size-fits-all-treatment/standard care intervention) to improve physical activity and dietary behavior in employees. Effects on single health behaviors, psychological predictors of behavior change (intention, planning, and social support), BMI, and lifestyle (multiple behavior index combining physical activity and nutrition) were examined (see hypotheses 1-3). In addition, this study explored whether characteristics of the workplace, that is, whether or not the workplace was physically demanding, moderated the efficacy of the intervention (explorative analysis). The second aim of the study was to examine why the intervention was effective and to identify the psychological variables that may account for changes in behaviors (hypothesis 4). The aforementioned hypotheses are detailed in the following section.

The main intervention effects (contrasting the stage-matched intervention to an active control condition) were hypothesized in terms of (1) more behavioral change in physical activity and dietary behavior (single behavior indicators, hypothesis 1a) and adoption of a healthy lifestyle (the synthesis of both behaviors, hypothesis 1b). We also expected improvements in (2) psychological predictors of behavior change (intention, planning, and social support, hypothesis 2) and (3) BMI (hypothesis 3). Finally, we expected (4) that those individuals who successfully had increased intention, planning, and social support due to the intervention would be more likely to adopt a healthy lifestyle (mediation effect, hypothesis 4).

Methods

Participants and Procedure

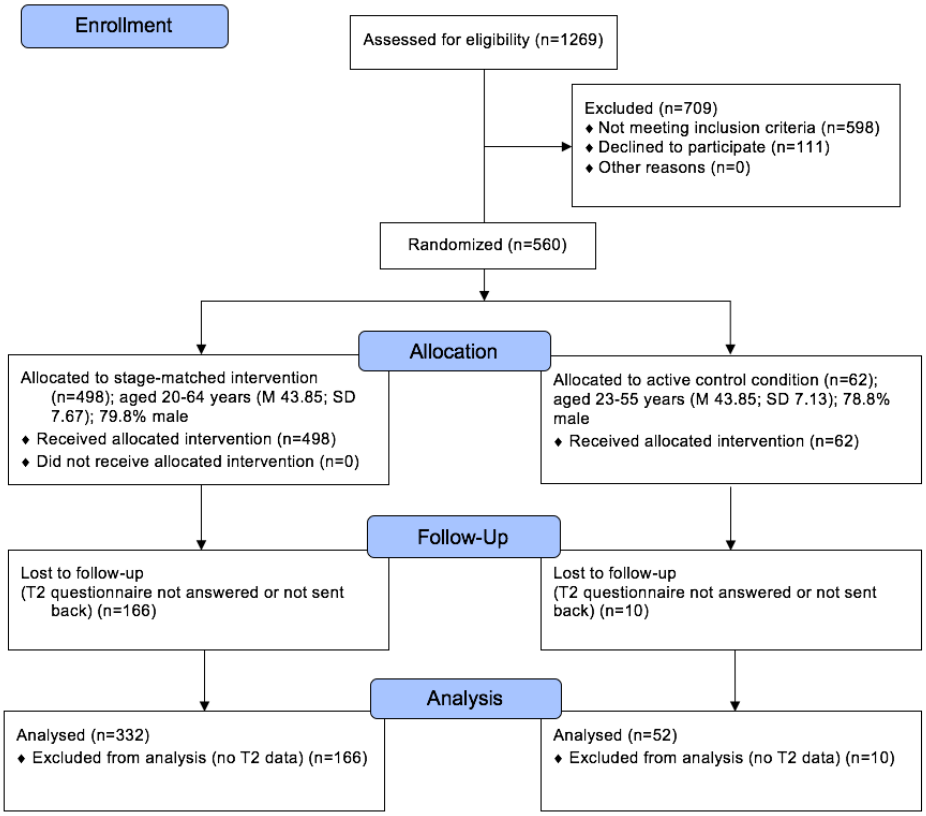
A total of 1269 shiftworkers in more or less physically demanding positions (eg, train drivers, ticket inspectors, track workers) were recruited during a routine medical examination which takes place once in every 3 years (Figure 1). Posters were put on the wall in the entrance of the company's physician office to make employees aware of the study and to increase their willingness to participate in it. All employees were asked face-to-face by a trained research assistant to participate in the study while waiting to see the company's physician.

If they agreed and signed an informed consent form, they were introduced to the computer kiosk with the computer-based, closed survey and counseling intervention (ie, answers were automatically recorded by the online questionnaire). The consent form contained a study participants' personal code and his/her name plus the address to contact them again for the follow-up. Research assistants entered the personal code of the study participants into the system to register the employee and to give him/her access to the questionnaire and the intervention. This personal code was kept with the questionnaire entries to merge the data from the different measurement points later on. However, no names or address data were entered and absolute anonymity of the individual was ensured. Written consent forms containing information that directly identifies the participant (eg, name, address, date of birth) were kept in a locked place.

Different data security and quality measures were taken: unauthorized access was not possible because the baseline measurement (including the intervention) was only performed in the company and under the supervision of a research assistant. No personal data were recorded. Cookies were not used and Internet protocol check was not performed: 2 company-owned computers were used for participants to complete the survey and intervention; and because cookies are very dysfunctional they were not utilized for this study. The research assistant ensured that study participants were logged-in correctly with

their individual code and no one could observe this process. Anonymized electronic data were stored on secure servers.

Figure 1. Flowchart of participant progress through the study phases.



Ethics and Consent

All participants were informed about the purpose of the study (including information on the length of the questionnaire and data storage procedures) with a participant information form and an informed consent form. All procedures performed in this study were in accordance with the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study protocol was approved by the Deutsche Gesellschaft für Psychologie in Germany. Because this study was carried out in an occupational setting and approval was given by the works council including a confidentially note, no clinical trial registration was required.

The questionnaire before the intervention was mandatory to be filled in by every study participant to avoid unit-missings. However, whether or not study participants actually answered the individual questions was voluntary, except for 2 key questions that are mandatory: To match the intervention appropriately, the stage regarding nutrition and physical activity needed to be determined. Single-item missings for these 2 questions were prevented by a pop-up message asking participants to fill in the question as otherwise they could not get any further. With the exception of the 2 mandatory questions, no completeness check was performed. Study participants were able to review and change their answers (through a back button or the backspace button). In case of questions, the research assistants were at hand to reduce the risk of any drop-outs from the study.

Design

The research assistant also helped in case of any problems, for example, by asking study participants to get back to the questionnaire if they were distracted, reminding study participants to read each page carefully, providing some instructions if employees needed help with understanding the tasks or handling the computer keyboard or mouse. These specific measures were taken because pilot studies revealed that some older employees had insufficient computer literacy.

Further measures to control for potential atypical answering styles were not taken because the study participants should answer the questions as unbiased and spontaneously as possible. Irrespective of whether participants completed the questionnaire and the intervention, participants received a pedometer as incentive. Completeness of the data was only checked post hoc. Data were collected between October 2006 and June 2008 in Germany, and all materials outlined below were translated from German.

Inclusion criteria were not being diagnosed with diabetes, no acute myocardial infarction within the last year, no medical condition that conflicted with general recommendations for physical activity and fruits and vegetables consumption, and sufficient language competences. Eligible employees (N=384) taking part in both Time 1 (T1) and Time 2 (T2) were randomly assigned to either 1 of the 9 stage-matched intervention packages (see Figure 2 and description below, n=332), or the active control condition (n=52) by a computer algorithm with a likelihood of 1/10, using the software DynQuest [22]. Participants and research assistants were blinded to their

allocation for the duration of the study. The software in the background also managed the log file.

After providing informed consent, 560 participants completed the baseline T1 questionnaire on behavioral, psychological, and sociodemographic variables. Subsequently, the computer algorithm assigned participants either to the stage-matched intervention or to the active control condition. At T2, 1 month later, follow-up questionnaires were returned by 384 participants (completion rate, ie, users who finished the survey, was 68.6%), constituting the longitudinal sample that mostly included men (n=306, 79.7%). Participants were between 20 and 64 years of age, with a mean (men) age of 43.7 years (SD 7.6).

Figure 2. Experimental 9-group design for the stage-matched intervention. A: Intention formation for nutrition and physical activity; B1: intention formation for physical activity and plans for nutrition; B2: intention formation for nutrition and plans for physical activity; C1: intention formation for physical activity and relapse prevention for nutrition; C2: intention formation for nutrition and relapse prevention for physical activity; D: plans for nutrition and physical activity; E1: plans for physical activity and relapse prevention for nutrition; E2: plans for nutrition and relapse prevention for physical activity; F: relapse prevention for nutrition and physical activity. Numbers in brackets indicate the number of pages the particular intervention package consisted of.

Nutrition

Physical activity	Non-Intender	Intender	Actor
Non-Intender	A (34 pages)	B_1 (30 pages)	C_1 (32 pages)
Intender	B_2 (32 pages)	D (27 pages)	E_1 (29 pages)
Actor	C_2 (34 pages)	E_2 (29 pages)	F (31 pages)

Experimental Conditions

The stand-alone computer-based intervention consisted of a questionnaire, information about health authority guidelines for physical activity, and the respective experimental component. Trained research assistants helped in a nondirective manner in the event that questions arose. The stage-matched intervention consisted of 3 packages targeted at the 3 stages of the HAPA, namely, nonintenders, intenders, and actors, who were individually tailored (see Figure 2, for the content description see below). Tailoring consisted of using previous answers that were used later on for further tasks and questions or specific feedback if, for example, no answer was given or feedback on weight was to be given (details provided in the following section).

The staging algorithm first considered whether the behavior was already performed on a regular basis, separately for nutrition (ie, eating 5 portions of fruits and vegetables each day) and physical activity (ie, performing at least 30 minutes of volitional physical activity 3 times a week). If so, then these respondents were categorized as “actors.” If not, their level of behavioral intention was considered (eg, whether they intend to strive for the target behavior within the next month). If endorsed, respondents were categorized as “intenders”; if not, they were regarded as “nonintenders.”

Cross-tabulating the 3 subgroups for nutrition with the 3 subgroups for physical activity yielded a total of 9 cells (see Figure 2 for sample sizes within cells). In the following section, the intervention packages tailored for these 3 groups are described. Various behavior change techniques [23] were tailored to the characteristics of the participants. All interventions were used in this predefined format in terms of stage tailoring. The materials were developed based on previous intervention materials [19,24]. In addition, we ran focus groups and extensive pilot tests to ensure the usability and technical functionality of the material.

Intention Formation for Nonintenders

This package was specifically used for employees not intending to adopt the recommended behaviors. The intervention targeted risk awareness, outcome expectancies, goal setting, and self-efficacy. In the beginning, risk awareness was addressed by asking participants whether they led a rather inactive lifestyle or an active lifestyle and whether they ate high calories and fatty products or lots of fruit and vegetable. Participants were then informed about the connection between physical activity and diet and blood vessel fitness. They were then asked to rate whether they thought their blood vessels are rather clogged or in good shape. Both ratings had to be given on a visual analog scale, moving an indicator on the computer screen to the left (inactive lifestyle versus unhealthy diet, clogged vessels) or to the right (active lifestyle versus healthy diet, fit vessels).

The same method was used to address outcome expectancies: participants were asked to indicate how they would look if they would perform regular physical activity and eat fruits and vegetables instead of high-calorie and high-fat products; rather being obese (left side) or rather being normal weight (right side). Subsequently, the recommendation was provided that physicians as well as exercise and nutrition experts suggest exercising 3 times a week for at least 30 minutes, and eating at least five portions of fruits and vegetables daily. In addition, a statement was given, saying that this level of activity and nutrition is doable. Participants were then asked to think about the positive

consequences (pros) of meeting this behavioral goal. One example was given (Then I would feel better), and up to 4 fields were provided to fill in positive outcomes. Afterward, participants were asked to generate 1 potential negative outcome. One example was also given for this (Then this costs me a lot). If negative outcomes (cons) were stated, these were then contrasted with some pros by asking participants to come up with something that could balance the cons.

Individuals were then asked to set behavioral goals for the next 3 weeks. The instruction explicitly included setting small steps toward reaching the larger goal of becoming more physically active during leisure time. Examples were given such as "...go swimming after work" and "...add a tomato to my supper." The first given goal was then displayed again in the context of "I intend to..." and it was asked whether people were optimistic about attaining this goal, indicating "not very likely" (left side) or "rather likely" (right side) on a visual analog scale on the computer screen.

To address self-efficacy, the following instruction was given: "Become more confident! Think about how you could master attaining your goal on your own, and what could help you to become more active/eat more fruits and vegetables successfully? What would be your trip or trick?" With the last page, goal setting was addressed again by asking people to sum up, by checking the different options for becoming more physically active and eating healthier that they could concretely consider for themselves.

Planning for Intenders

This package was specifically intended for employees who have set a goal to change their behavior. The package included the generation of action plans and coping plans (for an overview of the evidence, see [25]). In the beginning, the general recommendations of the physicians as well as exercise and nutrition experts (ie, exercising 3 times/week for at least 30 minutes, and eating 5 portions of fruits and vegetables daily) were introduced to the participants to intensify goal setting. A statement was provided that this level of activity and nutrition is doable. In contrast to nonintenders, intenders were given the following information: it was explained that a day has 1440 minutes, and that 30 minutes/day could easily be allocated to volitional physical activity; and that one typically eats 3 meals and 1 or more snacks, which opens up ample opportunities to add or replace products by fruits and vegetables. Even in a busy day, it should be possible to exercise and eat healthily. Three examples were given to stimulate the participants' thinking toward different opportunities and different cues to action to actually facilitate behavior enactment, such as taking a sports bag to work or taking an apple for snack.

Participants were then asked to name up to 3 personal behavioral goals to meet the target of being physically active 3 times a week for 30 minutes or longer as well as to eat 5 portions of fruits and vegetables each day. These goals were then displayed on the next pages, always 1 goal on 1 slide, with the request that the participants generate an action plan (ie, to specify when, where, and how and, for activity only, how long to act).

To prompt formation of coping plans afterward, participants were also given an example of what could pose a risk to the maintenance of goals and their translation into action. An example for physical activity was bad weather that could prevent running in the park. For nutrition, an example was that no fresh fruit could be available while traveling. The instructions to formulate coping plans followed. The example of a suggested coping plan was doing some indoor activities such as swimming or visiting a fitness studio, or buying some fruit in a grocery store on the way to work. Subsequently, people were asked to identify up to 3 barriers to their own action plans. As with the goals, these barriers were then displayed on the next pages again, always 1 barrier on 1 slide, with the request that the participant generates a coping plan (ie, how to stick to the goal pursuit and find a different way to meet it).

Relapse Prevention for Actors

This package was specifically intended for employees already performing the goal behaviors. Action control and coping plans were addressed. Participants were asked to write down up to 3 experiences with their actions (to capture action control) in an identical format for the creation of action plans for intenders. However, instead of anticipating future situations, participants were asked to consider past situations. Individuals were asked to reflect on those actions and situations (showed on a respective page with the retrieved information), and on whether they would like to adjust aspects of them to maintain this behavior in the future. If the desire for change was expressed, individuals could record their new, adjusted action plan.

Afterward, participants received the coping and planning intervention. In this they were asked to generate up to 3 potential barriers to being active, and strategies on how to overcome these barriers (equivalent to the format for intenders).

Combination of the Different Packages for a Stage-Matched Intervention

The different packages were combined in the different stage-matched interventions displayed in Figure 2. In the beginning, a brief feedback was given on the former behavior and intention. Nonintenders were informed that their behavior did not meet the recommendations, and that they would work on strategies concerning how to change their behaviors with the following program. Intenders were congratulated on their decision to change their behavior, and also informed that the following program would assist them in doing so. Actors were congratulated for performing the target behavior. In addition, information on the difficulties in maintaining a former behavior was given, along with the fact that it is possible to prevent falling back into inactivity. They were informed that the following program would help them in developing such a maintenance strategy.

Some linking sentences were given between the package on physical activity (always first) and nutrition (always second). Such sentences were, for example, for nonintenders, "Wonderful! With this goal in mind the switches are on for a successful start with the change. Now clear the tracks for your first week goal. You determine the route."

In all packages except package A (in which people were nonintenders for both behaviors before the intervention, see Figure 2), a strategy training was also incorporated. With that, participants were challenged to reflect further on their anticipated barriers, and to think about what could be done about them in general. The instruction also included suggestions on the basis of best practice examples (stemming from pilot tests). A dummy variable reported, "If barriers crop up then..."

- "...I prioritize differently or come up with a completely different plan."
- "...I invest more energy in actually making things happen."
- "...I ask others to help me."
- "...I look for other people whom I could use as role models."

In the end, participants were given good wishes. In addition, they were instructed to identify options for rewarding themselves for approaching their goal, such as by buying oneself a flower (nonintenders). Intenders were cheered on for working so hard on their goals, and told that they should stick to their plans and start right away by performing them in practice. Actors also received positive feedback, along with the instruction to maintain their appropriate behavior. They were reminded to transfer coping plans into their daily life.

Active Control Condition (Standard Care Intervention)

The active control condition (one-size-fits-all-treatment) contained general health information, for example, on the etiology of obesity and the inter-relation between physical activity, nutrition, and energy expenditure. BMI was calculated by assessing the participants' weight and height, which they had previously entered in the questionnaire. Personalized feedback on participants' BMI was given, such as "You are overweight. If you have additional ailments, such as high blood sugar or problems with your joints or cardiovascular system, you should try to lose weight. Please contact your general practitioner!" Then a quiz on healthy dietary behavior was provided containing 13 questions on eating candy, rye products, milk and meat products or fast food, drinking soda beverages and alcoholic drinks, adding salt, and when to eat. Personalized feedback was given and a teaching session followed, giving educational information about the food pyramid including prompts on drinking and food preparation. Both the material assembling the one-size-fits-all intervention and the stage-matched intervention were developed based on focus groups outcomes, and pilot tests with the material were carried out to ensure the usability and technical functionality.

Measures

All questionnaire items stem from validated and well-tested measurement tools (eg, [20,24,26]). We also conducted pilot tests with the questionnaire items to ensure the usability and technical functionality especially with its electronic version. The items of the questionnaire were not randomized and all participants were asked to answer all items of the questionnaire with the exception of the report of number of children, which was only asked if employees indicated that they had children. In total, the questionnaire consisted of 70 questions.

Behavior (Single Behavior Indicators)

Physical activity was measured by an adaptation of the validated scale by Godin and Shephard [27]. Participants indicated how often per week and how long per session they performed strenuous physical activities that result in faster heart rate and excessive sweating (eg, intensive swimming) and moderate physical activities that are hardly exhausting with light sweating (eg, gymnastics). The total physical activity was the sum of sessions per week, multiplied by minutes per session. Development of physical activity behavior over time for the 2 different intervention groups and within the 3 stage groups is displayed in Figure 3.

Regarding nutrition behavior, participants were asked, "How many portions of fruits and vegetables did you eat per day?" The instruction was, "Please think about the last month. (Please note that potatoes do not count)." Two categories were provided, namely, "fruits" and "vegetables" [28]. The total portions of fruits and vegetables were the sum of the amounts reported each day. Development of nutrition behavior over time for the 2 different intervention groups and within the 3 stage groups is displayed in Figure 4.

Figure 3. Minutes of physical activity/week for the active control group (standard care, solid line) and the stage-matched group (dotted line) at T1 and T2.

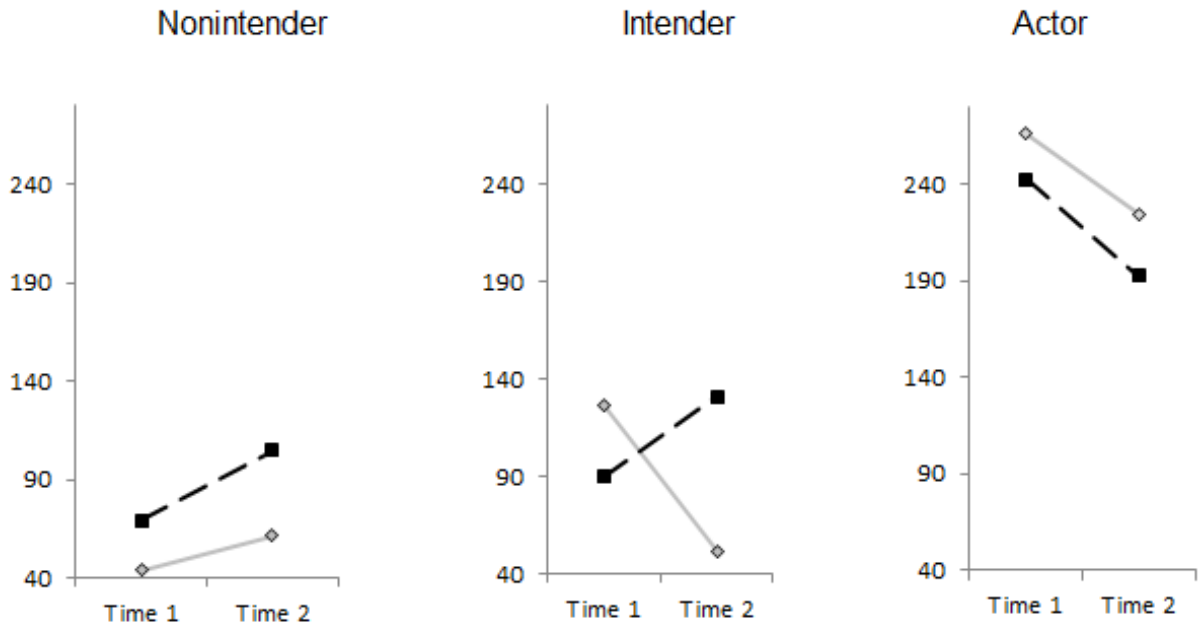
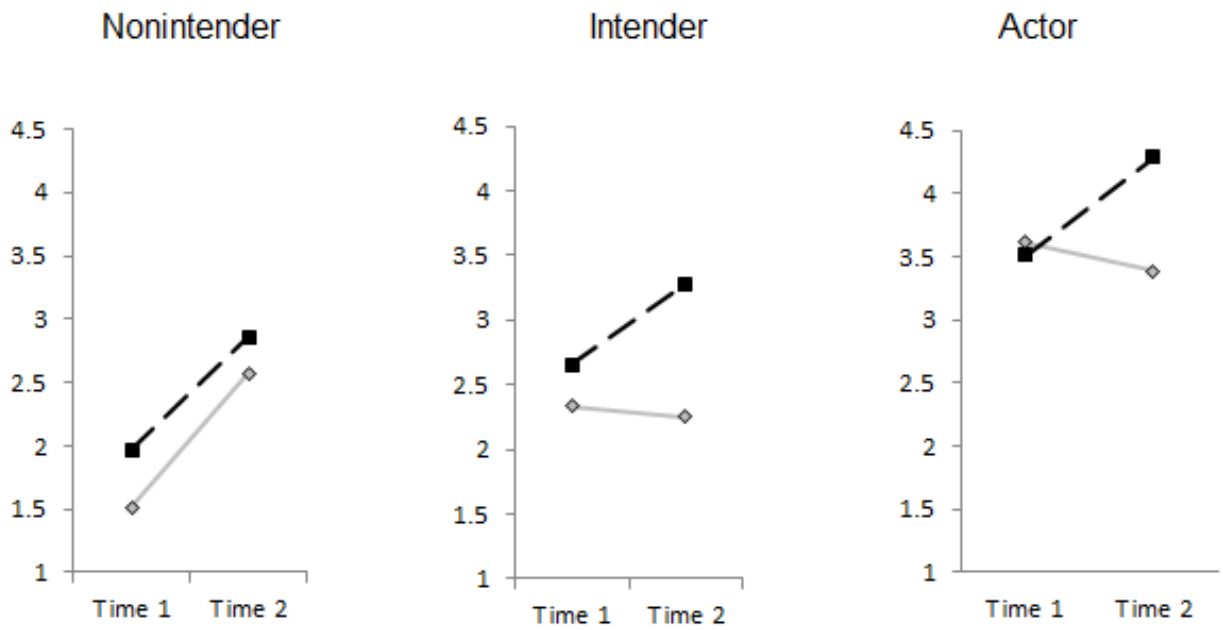


Figure 4. Portions of fruit and vegetables per day for the active control group (standard care, solid line) and the stage-matched group (dotted line) at T1 and T2.



Combined Healthy Lifestyle Indicator

To combine both behaviors, physical activity and healthy nutrition were categorized according to whether or not the participants met the recommendations. After piloting the usefulness of different criteria, the thresholds of 90-minute physical activity per week and 2 portions of fruits and vegetables were chosen. Both criteria were validated and both behaviors have been shown to be effective in improving health [29]. At T1, 46.6% of employees (179/384 eligible employees participating in T1 and T2) did not perform 90 or more minutes of physical activity per week, and 30.4% (117/384) did not eat 2 or more portions of fruits and vegetables per day. Both behaviors combined, 57.0% individuals (219/384) met only 1 or none of these 2 behavioral criteria and were categorized as having an unhealthy lifestyle at T1. As much as 42.9% (165/384) met both behavior recommendations and were categorized as exhibiting a healthy lifestyle at T1.

At T2, 33.9% employees (130/384 eligible employees participating in T1 and T2) did not perform 90 or more minutes of physical activity per week, and 17% of employees (66/384) did eat less than 2 portions of fruits and vegetables per day.

Both behaviors combined, 41.9% of individuals (161/384) met only 1 or none of the 2 behavioral criteria and were categorized as having an unhealthy lifestyle at T1. As much as 58.1% (223/384) of employees met both behavior recommendations and were categorized as exhibiting a healthy behavior at T2.

Psychological Predictors of Lifestyle Change: Intention, Planning, and Social Support

Physical activity intention was assessed with 2 items matching the 2 behavior intensities "I intend to perform the following activities at least 5 days per week for 30 minutes..." (1) "...strenuous (rapid heartbeats, sweating) physical activities" and (2) "...moderate (not exhausting, light perspiration) physical activities." The response options ranged from 1 to 4 "1=strongly disagree," "2=somewhat disagree," "3=somewhat agree," and "4=definitely agree." (4) The scale was aggregated, corresponding to the behavior measurement: strenuous and moderate activities correlated with r=.18 (P<.01) at T1, and with r=.23 (P<.01) at T2. Thus, items with discriminant validity were combined to obtain an index that reflects a broad construct.

Nutrition intention was also measured with regard to (1) fruits and (2) vegetables. The item was worded "I intend to..." "...eat 5 portions of fruits and vegetables a day" and "...eat fruits and vegetables with each meal." The answering options were as follows: "strongly disagree," "somewhat disagree," "somewhat agree," and "definitely agree." The 2 items were aggregated corresponding to the behavior measurement. The 2 items correlated with r=.53 (P<.01) at T1, and with r=.55 (P<.01) at T2.

Action planning was assessed with a single item based on procedures detailed in Lippke et al [24]. Activity-related plans were measured with the item "I have already planned exactly when, where, and how I want to be physically active." Nutrition-related plans were assessed by the item "I have already planned exactly when, where, and how I will eat 5 portions of fruits or vegetables throughout the day." The answering options were as follows: "strongly disagree," "somewhat disagree," "somewhat agree," and "definitely agree."

Social support was measured by answering the following 2 items: "How do you perceive your social environment?" (1) "My relatives are helping me to live healthily" and (2) "My friends and acquaintances are helping me to live healthily." The answering options were as follows: "strongly disagree," "somewhat disagree," "somewhat agree," and "definitely agree." The 2 items correlated with r=.57 (P<.01) at T1, and with r=.60 (P<.01) at T2, and were aggregated to a sum score.

Workplace Demands in Terms of Physical Activity

Study participants were instructed to think about the last weeks on their job and to rate whether they had performed physical activity for at least 30 minutes at work (eg, carry heavy stocks, walk long ways) for at least three to five times a week. Those agreeing to this item were categorized as working at a physically demanding workplace. Employees who indicated not performing this behavioral criterion were categorized as having a sedentary workplace.

These subjective ratings were validated with the reports of the employees' occupations in the company. Very few occupations were clearly categorized by all employees coherently such as "train inspectors" and "cleaning personnel" (with regard to a "physically demanding workplace," most other occupations such as "restaurant steward," "train driver," and "line manager" were rated differently by the respondents). Most occupations were rated by more employees as sedentary (eg, supervision was rated by 70.1%, 269/384 employees, as sedentary, train driver was rated by 65.5%, 251/384 employees, as sedentary, conductor was rated by 65.5%, 251/384 employees, as physically demanding, traffic controller was rated by 70.8%, 272/384 employees, as physically demanding). Thus, to acknowledge the individual situation at work, the rating of how physically demanding the work appeared to be for the individual was used to classify the workplace instead of categorizing for different occupations.

Sociodemographic Characteristics and Body Weight

Body height and body weight were used to calculate the BMI (determined by dividing weight in kilogram by squared height in meter) of all study participants. In addition, sex and age were assessed by self-report. Table 1 gives an overview on the descriptive statistics and intercorrelations of sociodemographics and lifestyle and physical demands at the workplace at T1 and T2.

Table 1. Descriptive statistics and correlations of study variables.

	Healthy lifestyle T1[a]	Healthy lifestyle T2[a]	Sex[b]	Age	Body mass index	Demanding workplace T1[c]	Demanding workplace T2[c]
Descriptives, n/N (%), ie, mean (SD)	165/384 (42.9)	223/384 (58.1)	77/384 (20.1)	43.69 (7.59)[d]	27.86 (4.86)[d]	163/384 (42.4)	181/384 (47.1)
Lifestyle T2	*r*=.36						
	P<.01						
Sex	*r*=.06	*r*=.10					
	P=.22	*P*=.06					
Age	*r*=-.05	*r*=-.12	*r*=-.06				
	P=.37	*P*=.02	*P*=.27				
Body mass index	*r*=-.06	*r*=-.12	*r*=-.01	*r*=.20			
	P=.34	*P*=.03	*P*=.81	*P*<.01			
Demanding Workplace T1	*r*=.01	*r*=-.08	*r*=-.05	*r*=.09	*r*=.05		
	P=.99	*P*=.11	*P*=.33	*P*=.09	*P*=.42		
Demanding Workplace T2	*r*=.10	*r*=-.05	*r*=-.05	*r*=.02	*r*=.07	*r*=.49	
	P=.05	*P*=.38	*P*=.33	*P*=.71	*P*=.21	*P*<.01	
	r=.02	*r*=.10	*r*=-.01	*r*=-.01	*r*=.01	*r*=.16	*r*=-.01
Intervention[e]	*P*=.69	*P*=.06	*P*=.84	*P*=.87	*P*=.92	*P*<.01	*P*=.92

[a]Lifestyle T1/T2 is an aggregate of both behavior recommendations (physical activity and eating fruits and vegetables).

[b]Sex: 0 indicates male (N=306); 1 indicates female (N=77); 1 employee did not indicate his/her sex.

[c]Demanding workplace T1/T2=0 indicates sedentary/not physically demanding; T1/T2=1 indicates physically demanding.

[d]Value presented as mean (SD)

[e]Intervention: 0 indicates active control condition; 1 indicates stage-matched intervention.

Analytical Procedure

Differential intervention effects on physical activity and nutrition (hypothesis 1a), psychological variables (intention, planning, and social support; hypothesis 2), and BMI (hypothesis 3) were tested by 2-factor repeated measures analysis of variance (ANOVA). The 2 factors were treatment (stage-matched intervention versus active control condition) and workplace (sedentary/physically not demanding versus physically demanding), and we examined their interaction with time as well as with each other and time.

Hypothesis 1b on the synthesis of the 2 behaviors was tested by employing frequency analyses (chi-square) and logistic regression (determining odds ratio, OR). Hypothesis 4 on the multiple mediator model was performed using an SPSS macro [30]. Residualized change scores were used, and confidence intervals were estimated by applying the bootstrap approach (5000 bootstrap resamples).

Results were reported based on the individuals participating in both measurement points. Imputed values were adopted for missing data within each measurement point in time using the expectation maximization algorithm in SPSS 22 (SPSS Inc, Chicago, IL, USA) [31]. However, this was only done if not more than 10% of items were missing, because otherwise the participation in the measurement point was interpreted as nonsufficient. All analyses were run with SPSS version 22. No methods to adjust for the representativeness of the sample were applied.

Results

Evaluation of Time, Treatment, and Workplace Demands on Single-Behavior Indicators and Psychological Predictors

Employees in the stage-matched intervention group (n=332) increased their physical activity in terms of minutes of strenuous and moderate exercise per week over time. The opposite effect was observed in individuals in the active control condition (n=52), in which employees decreased their mean activity over time (Figure 5A). However, neither the time nor the time × intervention nor the time × workplace × treatment effect was significant ($P \geq .15$; Table 2). Those employed in a sedentary workplace increased their activity from 144.71 (SD 187.28) minutes per week to 162.04 (SD 165.08) minutes per week. Study participants employed in a physically demanding workplace increased their activity from 169.71 (SD 240.38) minutes per week to 177.91 (SD 185.43) minutes per week (see Figure 5A for differential means). Thus, standard deviations were even larger than the means, which may have prevented the effects to be significantly different even though on a descriptive level they appeared distinct.

Table 2. Intervention efficacy evaluated in terms of changes over time tested in a 3-factorial repeated measures analysis of variance.

Test variable	Time	Time × intervention	Time × workplace	Time × workplace × intervention
Physical activity behavior	$F_{1,380}$=0.01	$F_{1,380}$=1.07	$F_{1,380}$=0.02	$F_{1,380}$=0.08
	η^2<.01	η^2<.01	η^2<.01	η^2<.01
	P=.47	*P*=.15	*P*=.44	*P*=.39
Physical activity intention	$F_{1,359}$=1.84	$F_{1,359}$=3.13	$F_{1,359}$=0.59	$F_{1,359}$=0.45
	η^2=.01	η^2=.01	η^2<.01	η^2<.01
	P=.09	*P*=.04	*P*=.23	*P*=.26
Physical activity planning	$F_{1,369}$=2.95	$F_{1,369}$=1.21	$F_{1,369}$=0.01	$F_{1,369}$=0.15
	η^2=.01	η^2<.01	η^2<.01	η^2<.01
	P=.04	*P*=.14	*P*=.47	*P*=.35
Nutrition behavior	$F_{1,380}$=17.92	$F_{1,380}$=0.55	$F_{1,380}$=0.37	$F_{1,380}$=0.26
	η^2=.05	η^2<.01	η^2<.01	η^2<.01
	P=.01	*P*=.23	*P*=.27	*P*=.32
Nutrition intention	$F_{1,377}$=2.58	$F_{1,377}$=4.03	$F_{1,377}$=0.12	$F_{1,377}$=0.03
	η^2=.01	η^2=.01	η^2<.01	η^2=.01
	P=.05	*P*=.03	*P*=.36	*P*=.43
Nutrition planning	$F_{1,375}$=1.83	$F_{1,375}$=4.45	$F_{1,375}$=0.04	$F_{1,375}$=0.01
	η^2=.01	η^2=.01	η^2<.01	η^2=.01
	P=.09	*P*=.02	*P*=.43	*P*=.50
General social support to live healthily	$F_{1,374}$=7.80	$F_{1,374}$=6.13	$F_{1,374}$=0.11	$F_{1,374}$=0.21
	η^2=.02	η^2=.02	η^2<.01	η^2<.01
	P<.01	*P*<.01	*P*=.37	*P*=.33
Body mass index	$F_{1,287}$=17.97	$F_{1,287}$=1.54	$F_{1,287}$=2.17	$F_{1,287}$=4.45
	η^2=.06	η^2=.01	η^2=.01	η^2=.02
	P<.01	*P*=.11	*P*=.07	*P*=.02

Regarding the portions of fruit and vegetable consumed, both groups increased their consumed portions per day. However, only the time factor was significant (*P*=.01), and not the intervention effect nor time × intervention nor the time × workplace × treatment effect (Table 2, Figure 5D). On average, employees consumed 2.45 (SD 1.66) portions per day at T1 and 3.22 (SD 1.71) portions per day at T2 (see Figure 5B for differential means).

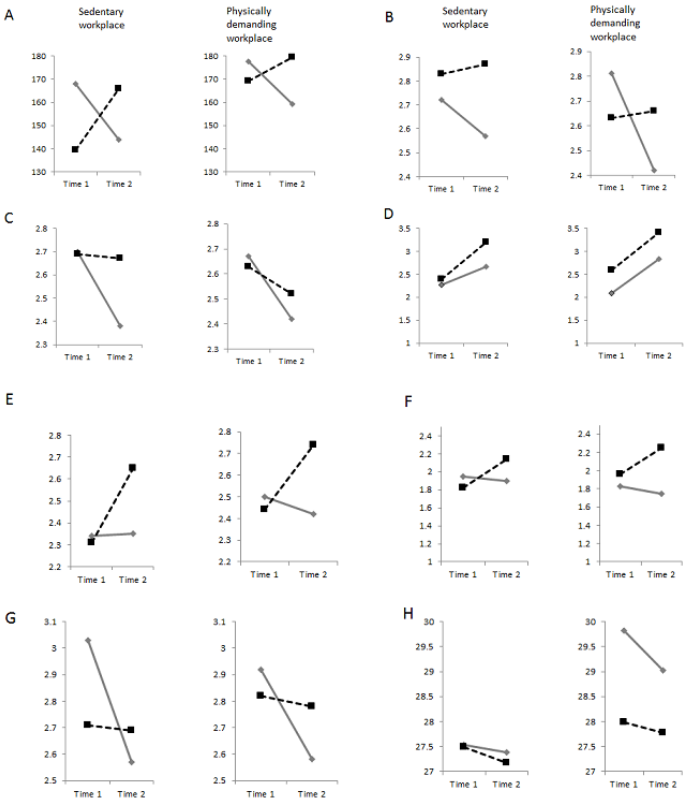
Effects were equally tested for intention and planning for each behavior domain as well as for social support and BMI. All means for the individuals in the active control condition versus the stage-matched intervention group are displayed in Figure 5, differentiated for study participants employed in a sedentary workplace and a physically demanding workplace.

Results from the repeated measures ANOVA are reported in Table 2. While descriptive changes in the measures were revealed over time and in favor of the stage-matched group (Figure 5), significant (*P*≤.04) time × treatment effects were only evident for physical activity intention, nutrition intention, nutrition planning, and social support (Table 2). An interaction of time × workplace or time × workplace × treatment could only be revealed for BMI. On average, study participants reduced their BMI from 27.75 kg/m^2 at T1 (86.05 kg) to 27.48 kg/m^{2}at T2 (85.23 kg).

This effect was about the same in the stage-matched group with a sedentary workplace (BMI_{T1}=27.49 kg/m^2; BMI_{T2}=27.18 kg/m^2) or with a physically demanding workplace (BMI_{T1}=27.98 kg/m^2; BMI_{T2}=27.77 kg/m^2). However, for the active control condition, those employees working in a sedentary workplace maintained their BMI over time (BMI_{T1}=27.54 kg/m^2; BMI_{T2}=27.38 kg/m^2). Those in the active control condition working in a physically demanding workplace started with a much higher BMI (BMI_{T1}=29.82 kg/m^2) and were able to reduce

their weight more than all other groups (BMI_{T2}=29.03 kg/m^2; Figure 5H).

Figure 5. Means for active-control group (standard care, solid line) and stage-matched group (dotted line) at T1 and T2. (A) Physical activity behavior (minutes/week). (B) Physical activity intention. (C) Physical activity planning. (D) Nutrition behavior (portions fruit and vegetables/day). (E) Nutrition intention. (F) Nutrition planning. (G) Social support. (H) Body mass index.



Evaluation of Time, Treatment, and Workplace Effects on Lifestyle Indicators

As the aim of the intervention was not only to change single behaviors but also to especially improve the employee's lifestyle consisting of 2 behaviors, changes in this combined outcome criterion were tested. Based on their nutrition and physical activity behavior, employees were categorized into whether or not they met the recommended criteria. The numbers and frequencies per group (differentiated by workplace: sedentary workplace versus physically demanding workplace) are shown for those already meeting or not meeting the recommendations at T1 in Table 3.

Descriptively, the stage-matched group outperformed in the active control condition for all subgroups. However, due to small sample sizes, this could only be tested for participants employed in a sedentary workplace (supporting the assumption of better effects of the stage-matched intervention) and the total group independently of the workplace (also in favor of the stage-matched group; Table 3). Significant support for the superiority of the stage-matched intervention over the standard-care intervention was only found for those study participants with a healthy lifestyle at T1 (P=.04). In this regard, for those with an unhealthy lifestyle at T1, the difference was not significant (P=.07) and was in favor of the stage-matched intervention. If all study participants were considered together, the standard-care treatment helped about 46% (24/52) of employees to practice a healthy lifestyle at T2, whereas the stage-matched intervention helped 59.9% (199/384) employees to practice a healthy lifestyle at T2 (ie, 13.7% more). This difference was statistically significant (χ^2_1=3.5; P=.04, for N=384).

Table 3. Performance of an unhealthy and a healthy lifestyle at T2, depending on T1 lifestyle and intervention.

Workplace	Lifestyle	Intervention	Lifestyle T2[a]			Statistic
			Unhealthy n (%)	Healthy n (%)	Total (100%)	
Sedentary workplace	Lifestyle at T1 unhealthy	Standard care	15 (60.0)	10 (40.0)	25	χ^2_1=0.1, *P*=.43[b]
		Stage matched	56 (55.4)	45 (44.6)	101	
	Lifestyle at T1 healthy	Standard care	5 (33.3)	10 (66.7)	15	χ^2_1=4.9, *P*=.04[b]
		Stage matched	9 (11.3)	71 (88.8)	80	
	Total	Standard care	20 (50.0)	20 (50.0)	40	χ^2_1=2.7, *P*=.07[b]
		Stage matched	65 (35.9)	116 (64.1)	181	
Physically demanding workplace	Lifestyle at T1 unhealthy	Standard care	6 (100.0)	0 (0.0)	6	Cannot be computed
		Stage matched	49 (56.3)	38 (43.7)	87	
	Lifestyle at T1 healthy	Standard care	2 (33.3)	4 (66.7)	6	Cannot be computed
		Stage matched	19 (29.7)	45 (70.3)	64	
	Total	Standard care	8 (66.7)	4 (33.3)	12	Cannot be computed
		Stage matched	68 (45.0)	83 (55.0)	151	
Total		Standard care	28 (53.8)	24 (46.2)	52	
		Stage matched	133 (40.1)	199 (59.9)	332	χ^2_1=3.5, *P*=.04[b]

[a]Lifestyle T1/T2=0 indicates not meeting both behavior recommendations (not performing ≥90 minutes of physical activity/week and not eating ≥2 portions of fruits and vegetables/day); T1/T2=1 indicates meeting both behavior recommendations (performing ≥ 90 minutes of physical activity/week and eating ≥ 2 portions of fruits and vegetables/day).

[b]N=384

Three models were tested with logistic regression analyses (Table 4). First, the performance of a healthy lifestyle T2 was predicted by sex, age, workplace demands, and BMI (all at T1). However, none of these 4 variables were a significant (*P*≥.07) predictor for a healthy lifestyle behavior at T2. Baseline lifestyle (T1) was included additionally as a predictor in model 2, which was related to a healthy lifestyle at follow-up: employees meeting the recommendations for a healthy lifestyle at baseline were 2 times more likely to also meet the recommendations at T2 (Table 4).

Table 4. Predicting follow-up lifestyle (T2).[a,b]

Variable	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Constant	2.58	7.61	4.82
Sex	1.45 (0.82-2.57), *P*=.20	1.43 (0.77-2.65), *P*=.25	1.44 (0.77-2.68), *P*=.25
Age	0.97 (0.94-1.00), *P*=.07	0.97 (0.94-1.00), *P*=.06	0.97 (0.94-1.00), *P*=.07
Demanding workplace T1	0.92 (0.75-1.12), *P*=.40	0.87 (0.70-1.07), *P*=.19	0.96 (0.91-1.01), *P*=.10
Body mass index	0.96 (0.91-1.01), *P*=.09	0.96 (0.91-1.02), *P*=.16	0.96 (0.91-1.01), *P*=.15
Lifestyle T1		2.26 (1.75-2.93), *P*<.01	2.25 (1.73-2.92), *P*<.01
Intervention [c]			1.96 (1.00-3.82), *P*=.05
R^2	.05, *P*=.03	.22, *P*<.01	.23, *P*<.01
ΔR^2		.17, *P*<.01	.01, *P*=.05

[a]Lifestyle T1/T2=0 indicates not meeting both behavior recommendations (not performing ≥90 minutes of physical activity/week and/or not eating ≥2 portions of fruits and vegetables/day); T1/T2=1 indicates meeting both behavior recommendations (perform ≥ 90 minutes of physical activity/week, and eating ≥ 2 portions of fruits and vegetables/day).

[b]Demanding workplace T1/T2=0 indicates sedentary/not physically demanding; T1/T2=1 indicates physically demanding.

[c]Intervention: 0 indicates active control condition; 1 indicates stage-matched intervention.

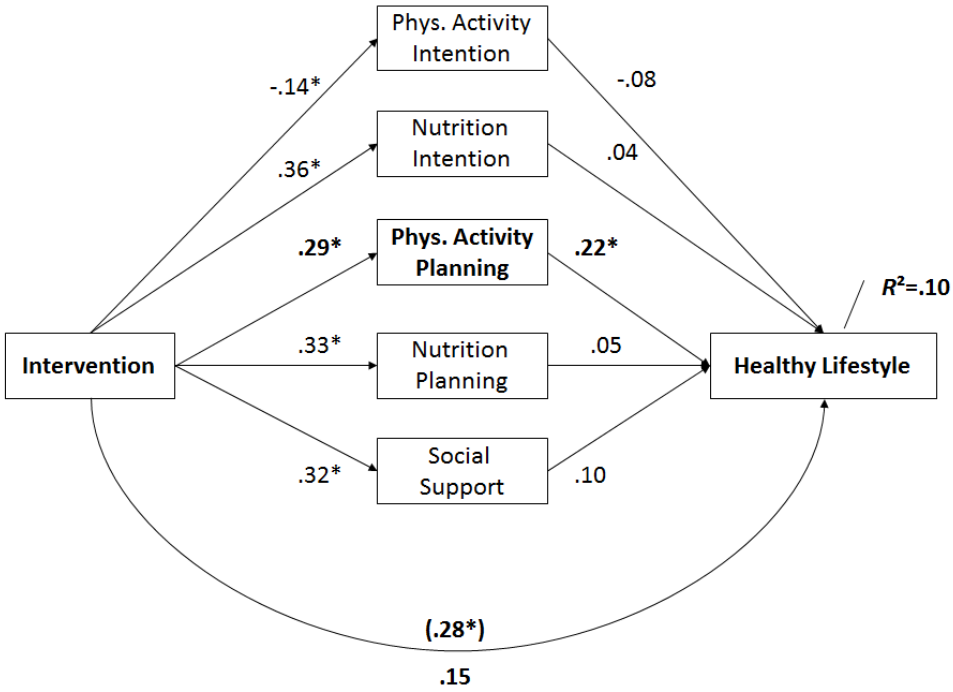
Likewise, in model 3, the treatment was tested in addition to the variables included in model 2. Receiving the stage-matched intervention in comparison with the active control condition was also a significant predictor (*P*=.05; Table 4) for validating

the previous results from frequency analysis (Table 3). Practically speaking, those employees in the stage-matched group were 2 times as likely to adopt or maintain a healthy lifestyle in comparison with those who received the active control condition. With model 3, almost one fourth of the variance within lifestyle at follow-up could be attributed to baseline lifestyle and treatment.

Testing Mechanisms of How the Treatment Facilitated a Healthy Lifestyle

Finally, a multiple mediator analysis [30] tested whether the effects of the intervention on lifestyle change (the synthesis of physical activity and nutrition) may be explained by changes in intention, planning, and social support (Figure 5). Residualized change scores obtained by regressing T2 scores on T1 scores were chosen for the putative mediators (Figure 6). Group assignment predicted changes in all social-cognitive variables, namely, in activity intention (beta=-.14, standard error, SE=.07; *P*=.04) and in nutrition intention (beta=.36, SE=.14; *P*=.01), and in activity planning (beta=.29, SE=.15; *P*=.05, shown in bold in Figure 6) and in nutrition planning (beta=.33, SE=.13; *P*=.01), as well as changes in social support (beta=.32, SE=.11; *P*=.01). Lifestyle change, as operationalized by meeting the recommendation toward physical activity and nutrition, was predicted only by changes in activity planning (beta=.22, SE=.05; *P*=.01, shown in bold in Figure 6) and by no other variable. After controlling for changes in these predictor variables, the relation between group assignment and behavior change was no longer significant (beta=.15, SE=.14; *P*=.28; without controlling: beta=.28, SE=.14; *P*=.05), which indicates that physical activity planning was a full mediator of the intervention effectiveness. The multiple mediator model accounted for 10% of the variance (R^2_{adj}=.08; *P*<.01) in lifestyle.

Figure 6. Mediation of the effect of the intervention on lifestyle changes by psychological variables. Significant changes are indicated by an asterisk.



Discussion

Preliminary Findings

This study aimed at gaining insights into computer-based health promotion for employees and, more specifically, the efficacy of a stage-matched intervention in comparison with an active control condition. Evaluated outcomes were behavior (primary outcome), intention, plans, social support, and lifestyle changes combining both behaviors and BMI (secondary outcomes). In addition, it was tested whether there was an effect of employees' workplace characteristics, that is, whether employees had high physical activity demands at work or not, on changes in primary and secondary outcomes. A total of 384 employees from a large logistics company took part in the study consisting of an in-house measurement point with an intervention component (online questionnaire and computer-based active control treatment/stage-matched intervention) and a mail-out questionnaire 4 weeks later.

Principal Results on Test of Hypotheses

The main expected intervention effects on single behavior and combined lifestyle were identified: in the stage-matched intervention group, significantly more study participants than in the active control group improved their lifestyle, operationalized as meeting the recommendations for physical activity and nutrition (hypothesis 1b). This is a practically important finding, as employees need to improve their lifestyles to improve their health as well as their risk of absenteeism and presenteeism. When evaluating single behavioral outcomes, the same trend was observed: in comparison with individuals in the active control group, individuals in the stage-matched intervention group reported less decrease of physical activity as well as more increase of consumption of fruits and vegetables

over time (hypothesis 1a). Although differential intervention effects surfaced on a descriptive level, we could detect a significant time × treatment interaction only for selected variables and conditions. Nonsignificant effects should be interpreted with keeping in mind that standard deviations of behavior were very high and even larger than the means of behavior. Overall, our findings on multiple behavior change replicate previous studies, which showed effects on both behaviors [13,10]. Further testing the practical importance of the intervention on lifestyle revealed that employees receiving the stage-matched intervention were 2 times more likely to adopt a healthy lifestyle (the synthesis of nutrition and physical activity) than employees in the active control group. Taking the different findings together, hypothesis 1 was partially confirmed.

Expected intervention effects on psychological predictors of behavior (change) were also identified: when testing the effects of time × intervention on the 5 psychosocial test variables (intention and planning per behavior and general social support), 4 were found to be significant. While the stage-matched intervention prevented the naturally occurring decline in physical-activity-related cognitions and social support, it was even able to increase the cognitions in the nutrition domain in comparison with the active control group. This confirmed hypothesis 2 with a majority (4/5) of the tested variables.

When testing the hypothesis on the intervention effects on BMI, we found an unexpected effect in terms of an interaction with workplace (ie, workplace × intervention group × time): whereas individuals in the stage-matched group in a sedentary workplace decreased their BMI more strongly than employees assigned to the standard-care condition, the opposite effect was revealed in study participants in a physically demanding workplace. In other words, individuals in a physically demanding work environment showed a higher decrease in BMI if they had been allocated to the active control condition instead of the stage-matched intervention. The significant changes in BMI over time are in line with the assumption that computer-based behavior change interventions have a potential to facilitate prevention, which is coherent with the emerging literature [32]. In a Cochrane systematic review on interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people, it was found that such interventions significantly reduced body weight [15]. Our study is consistent with this finding and applies it to an occupational setting [2]. The finding that the active control condition was more successful in reducing BMI in employees working in physically demanding workplaces might be related to content of the active control treatment: it seems that explicitly addressing BMI and giving personalized feedback is especially effective for these at-risk individuals [13]. While the active control groups seemed to be advisable for addressing obesity topics, we have to conclude that hypothesis 3 was not supported. Thus, we only evaluated the mechanisms that translate the intervention effects on lifestyle changes and not on BMI.

The hypothesized changes in psychological predictors of behavior change (intention, planning, and social support) were found in the majority of the tested variables in the mediation analysis. However, physical activity planning appeared to be the only facilitator of the intervention efficacy: in the multiple mediation model, we found that individuals in the stage-matched intervention group who managed to maintain their physical activity plans (Figure 5C) are more likely to adopt or maintain a healthy lifestyle (Figure 6). This is especially remarkable as it shows the gateway effect of physical activity mechanisms on nutrition, which was found before [32]. In addition, our results suggest that generating action plans for physical activity can cross over to nutrition, whereas other motivational constructs such as intention appear to be rather behavior specific. Referring back to hypothesis 4, the data support the assumption of a mediator. However, only physical activity planning seems to operate as a mediator and not intention or social support. Thus, hypothesis 4 was only partially supported.

Results on the Interaction Between the Intervention, Time, and Workplace

The workplace characteristics emerged as a significant moderator for changes in BMI: for employees working in a sedentary workplace, the stage-matched intervention seemed to decrease their BMI more strongly than for employees in the active control condition. The opposite effect could be noted for employees working in a physically demanding workplace: here, individuals in the active control group started at a much higher level and seemed to decrease their BMI more than the stage-matched intervention group. However, this could also relate to methodological effects such as regression to the mean, and it is important to note that this interaction effect was small with only 1% of explained variance. In comparison with all other outcome variables, the effects of time on BMI were highest. However, the time effects accounted for only 6% of the variance in BMI change. In general, study participants decreased their weight over the time lag of 4 weeks (approximately 0.82 kg), which can be attributed to both computer-based treatments. Larger effects could be expected after a longer follow-up measurement point as bodily changes require more time.

In general, only 1 of 8 tests revealed a significant triple interaction with occupational physical activity (time × workplace × treatment). It remains unclear whether such an interaction is just too complex to explain additional variance in the other main factors and interactions. However, this might also direct toward the general merits of the computer-based intervention irrespective of workplace characteristics [14,15]. Findings on psychosocial predictors of the adoption of a healthy lifestyle also indicate that workplace, and age, sex, and BMI were not important in this process. Those employees who are engaged in a healthy lifestyle before were also more likely to maintain it. In more detail, previously active individuals were more than 2 times more likely to maintain a healthy lifestyle than those individuals who were not previously active.

Although employees in the stage-matched intervention were 2 times more likely to adopt a healthy lifestyle, this effect is mainly mediated by physical activity planning. Thus, it seems imperative to help people to plan their physical activity, which then not only helps them to become physically active as planned but also to eat more healthily. This matches previously detected gateway effects of physical activity on nutrition [32].

In general, the results are in line with previous studies, which were included in a recent review on health promotion interventions implemented by occupational health services [13]. Coherent with findings on computer-based interventions, significant effects were also found with regard to dietary change and physical activity [14]. The applied stage-matching approach is a very parsimonious option to allocate participants to intervention packages. Although the allocation is based on a single item (ie, stage algorithm), the intervention packages cover a number of key psychological constructs that are assumed to be important. Alternatively, participants could receive interventions based on their answers to each individual, psychological construct (ie, construct tailoring). This would require much more complex algorithms for tailoring the intervention to the needs of the recipients [14,17]. Overall, we were able to demonstrate the advantage of the stage-matched intervention over the one-size-fits-all intervention (ie, active control group). However, this study might also be seen as showing how important it is to include the matching of key constructs in an intervention in general.

Future Directions

Overall, our findings support the usefulness of stage-matched interventions. However, it remains unclear whether participants in the active control group would have benefitted more if they had received not only an information-based educational treatment but also a complex one that included more powerful constructs for behavior change such as self-efficacy, planning, and action control. Thus, the gains that were observed for parts of the entire sample might have also occurred in different subsamples if the same treatment components were provided. To examine this further, fully balanced match-mismatch research designs are needed [33].

Limitations

Some limitations need to be mentioned. A selection bias of study participants might be possible due to the following factors: the context of the company's physician office (eg, employees might have expected more advices than if the kiosk would have been in a cafeteria of the company), the open disclosure of the study aims, and the posters with the prompt to stay fit (consequently, more motivated people might have agreed to participate), compared with using a bogus story for alternatively recruiting study participants. Thus, these limitations should be taken into account when interpreting the results.

In addition, the current data are based on online self-reports. Online studies give researchers the potential to reach large samples of persons with diverse socioeconomic status and age, and from different geographic regions [22]. Although the validity of self-reports on health behaviors appears to be satisfactory and the utilized assessment was previously validated [27], further validity studies of (online) self-reports should replicate and extend the results of this study. Furthermore, only short-term effects were investigated. Long-term effects may be studied in greater depth in the future.

Thus, while the study had limitations (eg, self-report measures, single intervention session, short follow-up measurement point), future studies should test the findings using more extended follow-ups and recording objective behavioral outcomes. Moreover, in this study the efficacy was evaluated only in terms of self-reported behavioral data and social-cognitive predictors of behavior change. Usability testing employing eye-tracking technology could add to this in the future, as this has been shown to be an important facet of evaluation research in natural settings [34].

Conclusion

To conclude, for the practice of occupational health promotion, parsimonious computer-based interventions on multiple health behaviors open avenues for reaching more employees, especially those who are "on the road" as part of their job and may not have access to company-owned, on-site support programs (eg, face-to-face counseling). Upscaling individual-level, multiple behavior workplace health promotion programs is a key to preventing and managing chronic diseases. This is especially imperative among the workforce due to the high proportions of the total cost of productivity loss due to sick leave and disability pensions attributable to obesity and obesity-related diseases [7].

Such interventions can be implemented either as an independent, stand-alone program or as a supplement to existing on-site offers (eg, counseling). Independent, computer-based programs might be particularly appealing to shift or remote workers who do not have access to trained in-person counselors. Many employees could, thereby, be helped to be active and to stay healthy. Theoretical implications could be to further include the human factor mechanisms that translate intervention effects into lifestyle changes. Planning, as a central variable, should especially be considered further in occupational and organizational health promotion. In addition, transfer effects, from 1 behavior domain to another [26,32], should be researched in more depth to explore synergetic effects.

Acknowledgments

The authors thank Kureva Pritchard for his editorial work. The work, which lead to this paper, was partially funded by a research grant from the German Ministry of Education and Research (Bundesministeriums für Bildung und Forschung, BMBF, within the framework "Engagement-Learning-Competence Development: Innovation for a Modern Working World"—"Arbeiten-Lernen-Kompetenzen entwickeln. Innovationsfähigkeit in einer modernen Arbeitswelt"—Grant No 01HH12002).

Authors' Contributions

SL, AW, and RS contributed to the conception and design of the study, including the intervention. SL, LF, and AW performed data collection, and SL performed the statistical analyses. LF and RS provided guidance to the presentation of results. All authors were involved in the interpretation of the data and in writing the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bilger M, Finkelstein EA, Kruger E, Tate DF, Linnan LA. The effect of weight loss on health, productivity, and medical expenditures among overweight employees. Med Care 2013 Jun;51(6):471-477 [FREE Full text] [doi: 10.1097/MLR.0b013e318286e437] [Medline: 23632594]
2. Cook RF, Hersch RK, Schlossberg D, Leaf SL. A Web-based health promotion program for older workers: Randomized controlled trial. J Med Internet Res 2015 Mar;17(3):e82 [FREE Full text] [doi: 10.2196/jmir.3399] [Medline: 25830503]
3. Demerouti E, Le Blanc PM, Bakker AB, Schaufeli WB, Hox J. Present but sick: A three - wave study on job demands, presenteeism and burnout. Career Dev Int 2009 Feb 20;14(1):50-68. [doi: 10.1108/13620430910933574]
4. van Scheppingen AR, de Vroome EM, ten Have KC, Zwetsloot GI, Bos EH, van Mechelen W. Motivations for health and their associations with lifestyle, work style, health, vitality, and employee productivity. J Occup Environ Med 2014 May;56(5):540-546. [doi: 10.1097/JOM.0000000000000143] [Medline: 24806567]
5. Lehnert T, Stuhldreher N, Streltchenia P, Riedel-Heller SG, König H. Sick leave days and costs associated with overweight and obesity in Germany. J Occup Environ Med 2014 Jan;56(1):20-27. [doi: 10.1097/JOM.0000000000000065] [Medline: 24351899]
6. Janssen I, Katzmarzyk PT, Srinivasan SR, Chen W, Malina RM, Bouchard C, et al. Utility of childhood BMI in the prediction of adulthood disease: Comparison of national and international references. Obes Res 2005 Jun;13(6):1106-1115. [doi: 10.1038/oby.2005.129] [Medline: 15976154]
7. van Duijvenbode DC, Hoozemans MJ, van Poppel MN, Proper KI. The relationship between overweight and obesity, and sick leave: A systematic review. Int J Obes (Lond) 2009 Aug;33(8):807-816. [doi: 10.1038/ijo.2009.121] [Medline: 19528969]
8. Gortmaker SL, Swinburn BA, Levy D, Carter R, Mabry PL, Finegood DT, et al. Changing the future of obesity: Science, policy, and action. Lancet 2011 Aug;378(9793):838-847. [doi: 10.1016/S0140-6736(11)60815-5]
9. Birdee GS, Byrne DW, McGown PW, Rothman RL, Rolando LA, Holmes MC, et al. Relationship between physical inactivity and health characteristics among participants in an employee-wellness program. J Occup Environ Med 2013 May;55(5):514-519 [FREE Full text] [doi: 10.1097/JOM.0b013e31827f37d7] [Medline: 23618884]
10. Mastellos N, Gunn LH, Felix LM, Car J, Majeed A. Transtheoretical model stages of change for dietary and physical exercise modification in weight loss management for overweight and obese adults. Cochrane Database Syst Rev 2014;2:CD008066. [doi: 10.1002/14651858.CD008066.pub3] [Medline: 24500864]
11. Rongen A, Robroek SJ, van Lenthe FJ, Burdorf A. Workplace health promotion: A meta-analysis of effectiveness. Am J Prev Med 2013 Apr;44(4):406-415. [doi: 10.1016/j.amepre.2012.12.007] [Medline: 23498108]
12. Bartlett YK, Sheeran P, Hawley MS. Effective behaviour change techniques in smoking cessation interventions for people with chronic obstructive pulmonary disease: A meta-analysis. Br J Health Psychol 2014 Feb;19(1):181-203 [FREE Full text] [doi: 10.1111/bjhp.12071] [Medline: 24397814]
13. Kwak L, Hagströmer M, Jensen I, Karlsson ML, Alipour A, Elinder LS. Promoting physical activity and healthy dietary behavior: The role of the occupational health services: A scoping review. J Occup Environ Med 2014 Jan;56(1):35-46. [doi: 10.1097/JOM.0000000000000012] [Medline: 24351888]
14. Noar SM, Harrington GN, Van Stee SK, Aldrich SR. Tailored health communication to change lifestyle behaviors. Am J Lifestyle Med 2010 Nov 19;5(2):112-122. [doi: 10.1177/1559827610387255]
15. Wieland LS, Falzon L, Sciamanna CN, Trudeau KJ, Brodney S, Schwartz JE, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. Cochrane Database Syst Rev 2012;8:CD007675 [FREE Full text] [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]
16. 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]
17. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010 Sep;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
18. Craciun C, Schüz N, Lippke S, Schwarzer R. Enhancing planning strategies for sunscreen use at different stages of change. Health Educ Res 2012 Oct;27(5):857-867 [FREE Full text] [doi: 10.1093/her/cys091] [Medline: 22907532]
19. Schwarzer R, Cao DS, Lippke S. Stage-matched minimal interventions to enhance physical activity in Chinese adolescents. J Adolesc Health 2010 Dec;47(6):533-539. [doi: 10.1016/j.jadohealth.2010.03.015] [Medline: 21094429]

20. 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]
21. Wiedemann AU, Lippke S, Reuter T, Ziegelmann JP, Schwarzer R. How planning facilitates behaviour change: Additive and interactive effects of a randomized controlled trial. Eur J Soc Psychol 2011 Jan 19;41(1):42-51. [doi: 10.1002/ejsp.724]
22. Rademacher JD, Lippke S. Dynamic online surveys and experiments with the free open-source software dynQuest. Behav Res Methods 2007 Aug;39(3):415-426. [doi: 10.3758/BF03193011]
23. Michie S, Johnston M. Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. Health Psychol Rev 2012 Mar 15;6(1):1-6 [FREE Full text] [doi: 10.1080/17437199.2012.654964]
24. 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]
25. 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]
26. Fleig L, Küper C, Lippke S, Schwarzer R, Wiedemann AU. Cross-behavior associations and multiple health behavior change: A longitudinal study on physical activity and fruit and vegetable intake. J Health Psychol 2015 May;20(5):525-534. [doi: 10.1177/1359105315574951] [Medline: 25903240]
27. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985 Sep;10(3):141-146. [Medline: 4053261]
28. Schwarzer R, Schuz B, Ziegelmann JP, Lippke S, Luszczynska A, Scholz U. Adoption and maintenance of four health behaviors: Theory-guided longitudinal studies on dental flossing, seat belt use, dietary behavior, and physical activity. Ann Behav Med 2007 Apr;33(2):156-166. [doi: 10.1080/08836610701308221] [Medline: 17447868]
29. Annesi JJ, Mareno N. Temporal aspects of psychosocial predictors of increased fruit and vegetable intake in adults with severe obesity: Mediation by physical activity. J Community Health 2014 Jun;39(3):454-463. [doi: 10.1007/s10900-014-9828-8] [Medline: 24481711]
30. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods 2008 Aug;40(3):879-891. [Medline: 18697684]
31. Enders CK. A primer on maximum likelihood algorithms available for use with missing data. Struct Equ Modeling 2001 Jan 1;8(1):128-141 [FREE Full text] [doi: 10.1207/S15328007SEM0801_7]
32. Fleig L, Kerschreiter R, Schwarzer R, Pomp S, Lippke S. 'Sticking to a healthy diet is easier for me when I exercise regularly': Cognitive transfer between physical exercise and healthy nutrition. Psychol Health 2014;29(12):1361-1372. [doi: 10.1080/08870446.2014.930146] [Medline: 24894668]
33. Godinho CA, Alvarez M, Lima ML, Schwarzer R. Health messages to promote fruit and vegetable consumption at different stages: A match-mismatch design. Psychol Health 2015 (forthcoming). [doi: 10.1080/08870446.2015.1054827] [Medline: 26010304]
34. Asan O, Yang Y. Using eye trackers for usability evaluation of health information technology: A systematic literature review. JMIR Human Factors 2015 Apr 14;2(1):e5 [FREE Full text] [doi: 10.2196/humanfactors.4062]

Abbreviations

ANOVA: analysis of variance
BMI: Body mass index
HAPA: health action process approach
OR: odds ratio
SE: standard error

Edited by G Eysenbach; submitted 30.03.15; peer-reviewed by AC Zarski, R Cook; comments to author 28.05.15; revised version received 07.07.15; accepted 24.07.15; published 01.10.15.

Please cite as:
Lippke S, Fleig L, Wiedemann AU, Schwarzer R
A Computerized Lifestyle Application to Promote Multiple Health Behaviors at the Workplace: Testing Its Behavioral and Psychological Effects
J Med Internet Res 2015;17(10):e225
URL: http://www.jmir.org/2015/10/e225/
doi:10.2196/jmir.4486
PMID:26429115

Original Paper

Translating Evidence Into Practice via Social Media: A Mixed-Methods Study

Stephen Maloney[1]; Jacqueline Tunnecliff[1]; Prue Morgan[1]; Jamie E Gaida[2,3]; Lyn Clearihan[1]; Sivalal Sadasivan[4]; David Davies[5]; Shankar Ganesh[6]; Patitapaban Mohanty[6]; John Weiner[7]; John Reynolds[7]; Dragan Ilic[7]

[1]Monash University, Frankston, Australia

[2]University of Canberra, The University of Canberra Research Institute for Sport and Exercise (UCRISE), Canberra, Australia

[3]University of Canberra, Discipline of Physiotherapy, Canberra, Australia

[4]Monash University, Sunway, Malaysia

[5]University of Warwick, Coventry, United Kingdom

[6]Swami Vivekanand National Institute of Rehabilitation Training and Research (SVNIRTAR), Cuttack, India

[7]Monash University, Melbourne, Australia

Corresponding Author:
Stephen Maloney
Monash University
PO Box 527
Frankston, 3199
Australia
Phone: 61 990 44240
Fax: 61 990 44812
Email: stephen.maloney@monash.edu

Abstract

Background: Approximately 80% of research evidence relevant to clinical practice never reaches the clinicians delivering patient care. A key barrier for the translation of evidence into practice is the limited time and skills clinicians have to find and appraise emerging evidence. Social media may provide a bridge between health researchers and health service providers.

Objective: The aim of this study was to determine the efficacy of social media as an educational medium to effectively translate emerging research evidence into clinical practice.

Methods: The study used a mixed-methods approach. Evidence-based practice points were delivered via social media platforms. The primary outcomes of attitude, knowledge, and behavior change were assessed using a preintervention/postintervention evaluation, with qualitative data gathered to contextualize the findings.

Results: Data were obtained from 317 clinicians from multiple health disciplines, predominantly from the United Kingdom, Australia, the United States, India, and Malaysia. The participants reported an overall improvement in attitudes toward social media for professional development (P<.001). The knowledge evaluation demonstrated a significant increase in knowledge after the training (P<.001). The majority of respondents (136/194, 70.1%) indicated that the education they had received via social media had changed the way they practice, or intended to practice. Similarly, a large proportion of respondents (135/193, 69.9%) indicated that the education they had received via social media had increased their use of research evidence within their clinical practice.

Conclusions: Social media may be an effective educational medium for improving knowledge of health professionals, fostering their use of research evidence, and changing their clinical behaviors by translating new research evidence into clinical practice.

(J Med Internet Res 2015;17(10):e242) doi:10.2196/jmir.4763

KEYWORDS

social media; medical informatics; evidence-based practice; e-learning

Introduction

Greater accessibility to the Internet and mobile technologies has seen unprecedented opportunities for populations to access health information, and for health professionals to access new knowledge. In 2012, Facebook reported that it was hosting over one billion users, or 15% of the world's population [1] and Twitter announced its users were sending 500 million tweets per day [2]. The possibilities this presents in health care for increasing the movement of knowledge from "bench to bedside" is however still relatively unexplored. It is currently estimated that 80% of research evidence relevant to clinical practice never reaches clinicians delivering patient care [3]. One of the key barriers for the translation of evidence from demonstrated benefit into practice is the limited time available for clinicians to search for, and appraise, emerging evidence [4]. This raises the issue of whether social media could provide a bridge between health researchers and health service providers, allowing researchers to directly message peer-reviewed findings to clinicians, and for clinicians to receive the messages on their Web-enabled devices at the point of care [5,6].

In a health care context, one definition of social media is a "collection of Web-based technologies that share a user-focused approach to design and functionality, where users can actively participate in content creation and editing through open collaboration between members of communities of practice" [7]. Social media, by overcoming barriers to engagement associated with geography, can expand community-of-practice networks, connecting practitioners irrespective of whether they are rural, remote, or metropolitan based. The global reach of the Internet facilitates a sharing and evolution of national and global perspectives. Leading global and national health authorities, journals, research centers, health professional education societies, universities, and even hospitals now have an active social media presence, which indicates the increasing recognition of the importance of a social media presence among leaders and governing bodies of health professional groups. Bergl et al have demonstrated the use of Twitter feeds to advance an internal medicine residency program's educational mission, where Twitter enhanced the residents' overall education in residency [8]. The Cochrane Collaboration registered its first systematic review title on social media for clinical excellence in 2014. Arguably, social media has developed a "formal," professional side to its personality.

Although intuitively there seems to be mutual benefits to be gained from connecting clinicians directly with researchers, the issue has always been how do we create and facilitate effective connections? Social media would appear to offer this opportunity, as it becomes an acceptable medium between health researchers and clinicians in translating emerging evidence into practice [9]. However, in spite of growing research related to the pedagogical value of the media from an educational perspective and the role of blended learning modalities versus traditional formats for knowledge translation [10-14], the literature on the impact of social media on education has largely centered on opinions and arguments around culture change and aspects of the learning experience [7,15-17]. A systematic review exploring the uses, benefits, and limitations of social media for health communication was conducted by Moorhead et al [15]. Their review identified three overarching benefits of social media: increased interaction with others, more available shared and tailored information, and increased accessibility and widening access to health information. However, the vast majority of studies identified within their review were of a poor quality, largely due to the limitations of the methodologies used and the descriptive nature of the studies. Their review also identified the need to progress toward measuring the effectiveness of social media for knowledge translation. A further review conducted by Cheston et al exploring the use of social media in medical education identified only four studies that used an examination to assess knowledge change [7]. Of those studies, only one utilized a preintervention/postintervention design with data taken from the same participants, with empathy as their primary outcome of interest [18]. As yet, no published studies have reported an empirical evaluation of the impact and role of social media on translation of health research into practice.

We have thus undertaken a study that aims to determine the efficacy of social media as an educational medium to effectively translate emerging research evidence into clinical practice, through changing clinician knowledge and behaviors. The context for this trial was the management of tendinopathy. The scientific basis for managing tendon injuries has rapidly expanded in recent years [19], and it is also a condition that clinicians encounter regularly.

Methods

Design

The study used a mixed-methods approach. The primary outcomes of attitude, knowledge, and behavior change were assessed using a preintervention/postintervention, with qualitative data gathered to contextualize the quantitative results. Ethics approval for this research was obtained through the Monash University Human Research Ethics committee (CF 14/1372-2014000640).

Participants

The participant group was deliberately inclusive to reflect the nature and reach of social media platforms, and to authentically simulate health professional education via social media. The invitation to participate was distributed via social media—Monash Tendon Research Group social feeds—as well as via an email to the primary clinical affiliates (ie, partner health service sites) of (1) Monash University, Faculty of Medicine, Nursing and Health Sciences, Australia, (2) Monash University Malaysia, (3) Warwick Medical School, University of Warwick, United Kingdom, (4) University of Southern California, and (5) Swami Vivekanand National Institute of Rehabilitation Training and Research, India. Undergraduate students were eligible to participate if they were actively engaged in providing clinical care as part of workplace-based education.

Intervention

Participants took part in a short course, delivered through the social media platforms of Twitter and Facebook; these were

chosen as they are currently the two most widely used social media platforms. The course consisted of eight practice points provided to participants over 2 weeks, with half of the participants randomized to receive the education through Twitter, and half through Facebook. The practice points were developed by tendon researchers, in conjunction with educational and clinical experts. The practice points included "headline" information focused on the key finding or implication for practice, along with supporting materials that typically included a link to a peer-reviewed open access journal article or a podcast, of between 25 and 60 minutes duration, given by clinical experts. Only open access articles were used. Participants were encouraged to utilize the functionality of social media by responding to the practice points with their own comments and opinions, or forwarding messages of interest. Training in participation was provided in the form of a 5-minute video clip, explaining the details of creating an account, and in accessing and following the study's social media site. The participant would then be able to access the practice points during the course, interpret social media timelines, and contribute to the conversation. A detailed PDF file provided a written version for those who preferred this to video-based instructions.

Outcomes

Demographic data were collected about participants including country, gender, age, area of practice, years of experience, use of social media, and frequency of contact with people seeking health services for tendon conditions. Kirkpatrick's hierarchy of educational outcomes proposes that training effects should be examined for four levels of impact: (1) participant reaction, (2) knowledge, (3) change in behavior, and (4) change in health outcomes [20]. Outcomes across levels of impact 1-3 were studied, with the same measures taken at baseline and at 1 week after the intervention.

Participant reactions (Level 1) were measured using the Social Media Use and Perception Instrument (SMUPI). The SMUPI, a questionnaire validated in the population of internal medicine, seeks participant use and attitudes toward social media for continuing professional development [21]. The SMUPI consists of 10 items, each measured across a 5-point Likert scale, with the higher values representing more positive attitudes.

Knowledge (Level 2) was measured via an online examination administered 1 week before and after the short course. The examination included 16 questions; each practice point was allocated two examination questions. The first question related directly to the "headline" information of the practice point, and the second question required knowledge gained from the supporting information provided with the practice point. Questions were in multiple-choice format, with one correct answer and four distractors for each question (see Multimedia Appendix 1). The same questions were used for both baseline and posteducation examinations, except that the questions along with their answers and distractors were randomized to avoid answers based on pattern recognition. Self-rated knowledge was also measured with a 5-point Likert scale (1=very poor, 5=very good) asking the participant to self-rate their perceived knowledge of best practice in tendinopathy management. Self-rated confidence in interacting with people with tendinopathy was also measured using the same 5-point Likert scale.

Changes in behavior (Level 3) were determined by self-reported change in practice, or intended practice, after completing the program. Participants were first asked, "Has the education you have received via social media during this trial changed the way you practice, or intend to practice, with musculoskeletal clients?" followed by the open-text comment, "If you answered 'yes' to the question above, please indicate in the space provided below in what way the program has changed your management of tendon clients." Subsequently, participants had to respond to the following question: "Has the education you have received during this trial increased your use of research evidence within your clinical practice?"

Assessment at all levels of outcomes was administered via an anonymous online survey. Participants entered a unique identifying password that enabled the research team to match presurvey and postsurvey data for each participant to assess change across time. Participants received a certificate of course completion as an incentive, provided after the participant had completed the postcourse examination. To maintain participant anonymity, the mailing address provided for awarding the certificate was not linked to any survey data.

Data Analysis

Mixed linear models were used to analyze the repeated measurements of the participants. The analyses used all the available data and are based on the assumption that missing values are missing at random. No direct imputation of missing values was undertaken, with the mixed-models approach now regarded as standard [22]. The restricted maximum likelihood (REML) method, as implemented in the GenStat version 17 (VSN International) statistical package [23], was used to fit the models and calculate predicted means; *F* tests were used to test the main effects of time (pre and post). Pairwise least significant difference tests were based on these analyses and conducted at the 5% significance level. Diagnostic plots of residuals were checked to assess whether or not there were departures from the usual assumptions—homogeneity of variance and normality—required for optimal performance of these statistical tests. Analyses of the 5-point Likert scales also used the same approach as is customary with large datasets [24]. The analyses of binary response outcomes—"changed, or intention to change, practice" and "use of research evidence"—were based on logistic regression models, also fitted using GenStat version 17 (VSN International).

Results

Overview

Of the 317 total participants who participated in the study's data collection, 99 (31.2%) completed the baseline activities only, 45 (14.2%) completed the postbaseline activities only, and 173 (54.6%) completed the activities at both time points. A study flowchart is provided in Figure 1. Participants represented a spread of professions, ages, and countries, as summarized in Table 1.

The educational program was delivered as described. Screenshots to illustrate the presentation and format of the practice points developed and delivered via Twitter and Facebook are provided in Figures 2 and 3, respectively. The distribution of the eight practice points delivered were spread evenly over the 2-week period.

Table 1. Participant demographic data recorded at study baseline.

Descriptor		Participants (n=272), n (%)
Country		
	Australia	107 (39.3)
	India	28 (10.3)
	Malaysia	11 (4.0)
	United Kingdom	52 (19.1)
	United States	29 (10.7)
	Other	43 (15.8)
	Not recorded	2 (0.7)
Age in years		
	18-24	67 (24.9)
	25-34	123 (45.2)
	35-44	59 (21.7)
	45-54	16 (5.9)
	>54	7 (2.6)
Profession		
	Medicine	37 (13.6)
	Physiotherapy	193 (71.0)
	Podiatry	18 (6.6)
	Other	20 (7.4)
	Not specified	4 (1.5)

Figure 1. Study flowchart.

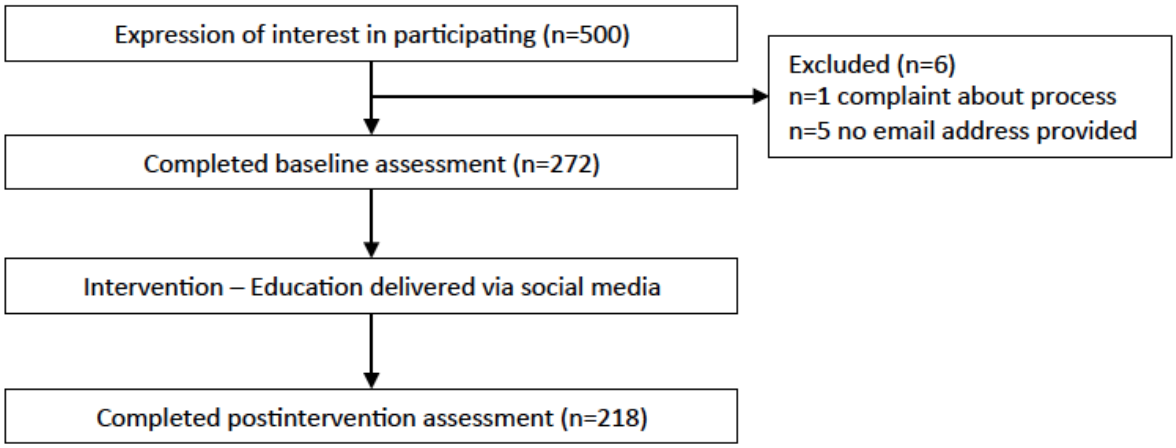


Figure 2. Screenshot of a practice point delivered via Twitter.

Figure 3. Screenshot of a practice point delivered via Facebook.

Evaluation Outcomes

Change in Participant Attitudes to Social Media and Participation (Kirkpatrick's Level 1)

The participants reported an overall improvement in attitudes toward social media for professional development using total scores for the SMUPI (premean=39.93, pre-SEM=0.41, postmean=41.32, post-SEM= 0.45; SED 0.44, 95% CI 0.57-2.29, *P*=.001), with improved attitudes reported for seven out of the 10 scale items (Table 2). A total of 8.7% (17/195) of participants reported actively engaging with fellow participants via social media during the study period.

Table 2. SMUPI[a] item predicted means from preanalysis/postanalysis.

SMUPI item	Preintervention			Postintervention			SED	95% CI	*P*
	n	Mean	SEM	n	Mean	SEM			
1. I would use social media to gain professional knowledge	269	4.075	0.051	213	4.226	0.056	0.058	0.036-0.266	*.009* [b]
2. I would use social media to enhance my education or professional development	269	4.084	0.050	212	4.250	0.055	0.057	0.054-0.277	*.003*
3. Social media would be useful for learning about professional development courses	269	4.231	0.048	211	4.268	0.053	0.058	-0.077 to 0.152	.53
4. I would be interested in social media for information about professional development opportunities	269	4.266	0.045	213	4.310	0.050	0.051	-0.057 to 0.145	.35
5. I would like to have professional development courses advertised to me by social media	269	4.047	0.050	212	4.207	0.056	0.060	0.042-0.279	*.007*
6. Professional development courses should use social media to enhance learning	268	3.950	0.052	213	4.095	0.058	0.063	0.020-0.270	*.02*
7. Social media is a professional way to assess professional development course content	269	3.550	0.056	212	3.894	0.063	0.067	0.213-0.476	*<.001*
8. Social media is an ethical way to engage professional development participants	269	3.674	0.053	213	3.972	0.059	0.061	0.178-0.418	*<.001*
9. Social media is an appropriate resource for professional development	269	3.820	0.053	212	4.056	0.058	0.060	0.119-0.354	*<.001*
10. Social media will be increasingly utilized for professional development in the future	268	4.277	0.048	211	4.327	0.054	0.057	-0.063 to 0.163	.42

[a]Social Media Use and Perception Instrument (SMUPI). The scale for all 10 items ranges from 1 (strongly disagree) to 5 (strongly agree).

[b]Statistically significant preanalysis/postanalysis changes are shown in italics.

Knowledge Change (Kirkpatrick's Level 2)

The participant responses to the self-rating of their knowledge of best practice in the management of tendon clients showed a significant increase in Likert scale means pretraining to post-training via the social media intervention (premean = 3.10, postmean = 3.65; difference = 0.55, 95% CI 0.44-0.65; $F_{1,215.7}$=103.2, P<.001). The participant responses to the self-rating of their confidence for interacting with people with tendinopathy also increased after the social media-based training (premean = 3.30, postmean = 3.71; difference = 0.41, 95% CI 0.31-0.52; $F_{1,211.8}$ = 59.32, P<.001).

The examination results demonstrated a significant increase in knowledge following the short course (premean = 44.52%, postmean = 61.70%; difference = 17.17%, 95% CI 14.06-20.28; $F_{1,201.4}$ = 121.8, P<.001). This shift in knowledge was made up of a significant increase in examination questions that were based on the "headline" information within the provided practice points (premean =3.97, postmean = 5.48; difference = 1.51, 95% CI 1.20-1.81; $F_{1,211.4}$ = 98.82, P<.001), as well as the examination questions based on the further activities (readings or podcasts) provided in the practice points (premean = 3.17, postmean = 4.42; difference = 1.25, 95% CI 0.98-1.53; $F_{1,206.7}$ = 80.75, P<.001). A subanalysis observed that the significant positive shift in knowledge occurred regardless of profession—medicine, physiotherapy, and podiatry—or country—the United Kingdom, Australia, the United States, Malaysia, and India. A summary table of the pretraining and post-training results for the knowledge change items is provided in Table 3.

Table 3. Predicted means for pretraining and post-training knowledge outcomes.

Item	Pretraining			Post-training			SED	95% CI	*P*
	n	Mean or test score	SEM	n	Mean or test score	SEM			
Self-rated knowledge of best practice, mean	270	3.104	0.051	217	3.649	0.056	0.054	0.438-0.651	<.001
Self-rated confidence, mean	271	3.298	0.053	217	3.714	0.058	0.055	0.308-0.524	<.001
Examination (test score), %	259	44.525	1.465	198	61.696	1.629	1.577	14.061-20.282	<.001

Behavior Change (Kirkpatrick's Level 3)

The majority of respondents (136/194, 70.1%) indicated that the education they had received via social media had changed the way they practice, or intended to practice. Those who answered in the affirmative were asked in what way their practice had changed. Thematic analysis of participant open-text responses revealed change in practice concentrated into three key themes: (1) use of evidence-based interventions, (2) patient monitoring, and (3) improved ability for shared (clinician and client) decision making. The following quotes have been provided below to contextualize these themes.

> *I am more willing to prescribe eccentric exercises and implement the knowledge I learned when treating future patients.* [Theme: Evidence-based interventions]
>
> *She [presenter in podcast] also mentioned using pain ratings after 24hrs post exercise to determine if the load was appropriate. I probably would have responded more to pain during the exercise as an indicator of proper exercise intensity.* [Theme: Patient monitoring]
>
> *The information provided on platelet rich plasma injections would allow me to provide patients with this information in a shared decision-making process.* [Theme: Shared decision making]

Similarly, a large proportion of respondents (135/193, 69.9%) indicated that the education they had received via social media had increased their use of research evidence within their clinical practice. Thematic analysis of open-text comments to better understand the participant responses revealed that the education provided a motivating reason as to *why* they should improve their evidence-seeking behavior. This largely occurred through surprise at how fast clinically relevant knowledge was advancing.

> *This course reminded me that even though I am only 2 years out of school, literature is rapidly advancing and changing; therefore it is vital to continue to seek literature to maintain evidence-based practice evidence.*

Participant use of research evidence was also influenced through the education by providing the participants with a new method in *how* to seek and access new information.

> *This was a great tool [social media] to catch up on the literature without having to pay for individual articles and not investing the time into untimely literature searches.*

Discussion

Principal Findings

This study evaluated the impact of delivering evidence-based knowledge to clinicians via social media. The findings contribute novel information to the field of medical education and evidence-based practice. Under our pretraining/post-training study conditions, clinicians reported positive changes in their attitudes and clinical behaviors, and demonstrated increased knowledge after their social media-based educational intervention. Although this is new empirical knowledge, it aligns with literature evaluating the effectiveness of blended learning approaches involving social media within health professional education [25], and researcher and clinician perspectives on the potential use of social media for effective professional development [10,16,21]. The key contrasting differences between earlier literature investigating knowledge change from social media [15,18,26] is the strength of the methodology and the context of the education focused on translating emerging research evidence into practice.

These preliminary findings, if validated in a randomized and controlled trial, would have implications across the health care, research, and education sectors. From the researcher's perspective, our results may provide new motivation to be outwardly engaged and more connected with clinicians at the coalface. The health researchers and health research centers who have cultivated a large social media following should be encouraged by the findings of this study, which support the potential effectiveness of their efforts. Researchers who are new to social media should be supported to actively disseminate research findings and participate in conversations with clinical stakeholders. From the clinician's perspective, social media is a legitimate form of continuing professional development, with the advantage of being flexible and asynchronous (ie, the material does not have to be accessed at the time of delivery). The extensive global use of social media by clinicians would enable access to minority areas of interest.

From the educator's perspective, social media provides an opportunity to link health professional learners to new learning opportunities, thereby expanding the classroom and providing a potentially international audience to locally produced education. Social media is most aligned with social-constructivist pedagogy, with the open seeking and sharing of information, encouraging collaboration and cooperation between learners. It may harness the educational advantage of fostering a sense of community for the learner: "Someone is contacting me, and I can talk to them. I can directly contact, and clarify issues with the leaders in the field; my perspectives can influence how experts consider the issues." This is a distinctly new model of student engagement, and a paradigm shift in how we conceptualize education. Education within social media enables a pedagogical approach that appropriately blends academic and personally elating experiences. More than just information seeking, learners have the chance to feel a part of the source of valued information, becoming cocreators and curators of the content. These influences are not limited to undergraduate education, or to the professional development of our existing health workforce industry. Industry requires graduates who have the skills and desire to critically appraise information, improve knowledge, and contribute to practice change [27]. Similarly, learners wish to enter a workforce that is adaptable as opposed to perpetuating outdated practices [17].

A key limitation to the effectiveness of social media for continued professional development, and the translation of new evidence into practice, is that social media relies heavily on the end user to be able to discern relevant information and ascertain

whether the information is based on sound evidence. Inability to critique the information presented may result in inaccurate information being created and perpetuated. Dangers also exist for the "education provider" in the social media pathway. Many researchers and clinicians who engage through social media platforms may not be aware of the safe and effective use of the medium, the etiquette of social media, or the application of relevant laws such as copyright and plagiarism. At present, there are limited training opportunities in the professional use of social media for health professionals.

Aside from limitations in the concept of social media for health professional education, limitations also exist within the study conducted that may affect the generalizability of the results. Knowledge change was assessed using the same 16-question multiple-choice exam at baseline and follow-up. This introduces the possibility of a learning effect. To counter this, both the examination questions and response items were presented in a random order. To minimize the risk of participant collaboration between participants during the examination, the learner was reminded that the examination was anonymous (ie, matched by code, rather than participant name) and that correct answers would be provided on completion of the study. There is also potential for reporting bias as participants were asked to self-assess the impact of the learning intervention. This limitation would be addressed in a randomized controlled trial of the intervention, which is recommended as a future line of investigation. A longer follow-up period may allow more detailed data to be collected on behavior change over time. The impact of self-selection and prior social media usage by participants has the potential to affect the participants' willingness to participate and their reaction to the educational platform.

This work identifies targets for research on social media within health professional education. Researchers, educators, and clinicians cannot ignore the reach of social media in health professional education. This reach is further evidenced by the logistics of this trial, a low-cost approach to bringing together a large number of clinicians from across the world—a research paradigm enabled by the nature of the intervention. Social media has developed into a convenient and acceptable link between clinicians and sources of information for professional development, as well as between researchers and clinicians. Rather than looking at whether or not social media is effective for health professional education, it may be time to look at how various modalities can be optimized, both in terms of how the messages are delivered and how learners can be supported to engage. In addition, the cost-effectiveness and sustainability of such models should be determined to allow informed risk management and to improve the adoption of such modalities. Further research may contribute to existing literature by focusing on the barriers and facilitators for leading researchers and clinicians to transition away from private social media forums to more public forums, allowing them to become more outwardly engaged in reducing the knowledge-to-practice gap in the health professions.

Conclusions

Social media appears to be an effective educational medium to provide information to clinicians for improving knowledge, fostering the use of research evidence by health professionals, and changing their clinical behaviors by translating new research evidence into clinical practice. Social media equalizes information sharing, thereby challenging historical hierarchies. It enables a direct information-sharing pathway and provides opportunities for discussion among clinicians, health researchers, professional associations, authors of journal articles, and other key industry stakeholders. Using social media to provide tailored and direct connections between researchers and clinicians facilitates the translation of new knowledge and practice, overcoming many of the obstacles within the "evidence pipeline." In this way, it can help reduce the evidence-to-practice gap—the essence of research translation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examination questions used in the assessment of knowledge change.

[PDF File (Adobe PDF File), 37KB - jmir_v17i10e242_app1.pdf]

References

1. Protalinski E. TNW News.: TNW News; 2013 Oct 30. Facebook passes 1.19 billion monthly active users, 874 million mobile users, and 728 million daily users URL: http://thenextweb.com/facebook/2013/10/30/facebook-passes-1-19-billion-monthly-active-users-874-million-mobile-users-728-million-daily-users/ [accessed 2015-10-13] [WebCite Cache ID 6cGDS0uMe]
2. Holt R. The Telegraph. London, UK: The Telegraph; 2013 Mar 21. Twitter in numbers URL: http://www.telegraph.co.uk/technology/twitter/9945505/Twitter-in-numbers.html [accessed 2015-10-16] [WebCite Cache ID 6cKQd2jVM]
3. Glasziou P. Evidence based medicine: Does it make a difference? Make it evidence informed practice with a little wisdom. BMJ 2005 Jan 8;330(7482):92 [FREE Full text] [doi: 10.1136/bmj.330.7482.92-a] [Medline: 15637375]
4. Lenfant C. Shattuck lecture--clinical research to clinical practice--lost in translation? N Engl J Med 2003 Aug 28;349(9):868-874. [doi: 10.1056/NEJMsa035507] [Medline: 12944573]

5. Butcher L. Oncologists using Twitter to advance cancer knowledge. Oncol Times 2010;32(1):8-10. [doi: 10.1097/01.COT.0000366137.51963.75]
6. Park M. Exploring healthcare opportunities in online social networks: Depressive moods of users captured in Twitter. In: Proceedings of Medicine 2.0: World Congress on Social Media, Mobile Apps, Internet/Web2.0. 2012 Presented at: Medicine 2.0: World Congress on Social Media, Mobile Apps, Internet/Web2.0; September 15-16, 2012; Boston, MA p. -5.
7. Cheston CC, Flickinger TE, Chisolm MS. Social media use in medical education: A systematic review. Acad Med 2013 Jun;88(6):893-901. [doi: 10.1097/ACM.0b013e31828ffc23] [Medline: 23619071]
8. Bergl PA, Narang A, Arora VM. Maintaining a Twitter feed to advance an internal medicine residency program's educational mission. J Med Internet Res 2015;1(2):1-6 [FREE Full text] [doi: 10.2196/mededu.4434]
9. Tunnecliff J. The acceptability among health researchers and clinicians of social media to translate research evidence to clinical practice: Mixed-methods survey and interview study. J Med Internet Res 2015;17(5):1-11.
10. Maloney S. The efficacy of Web-based versus face-to-face delivery of falls prevention exercise prescription education to health professionals: A randomised trial. J Med Internet Res 2011;13(4):1-10.
11. Shaw T, Long A, Chopra S, Kerfoot BP. Impact on clinical behavior of face-to-face continuing medical education blended with online spaced education: A randomized controlled trial. J Contin Educ Health Prof 2011;31(2):103-108. [doi: 10.1002/chp.20113] [Medline: 21671276]
12. Fordis M, King JE, Ballantyne CM, Jones PH, Schneider KH, Spann SJ, et al. Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: A randomized controlled trial. JAMA 2005 Sep 7;294(9):1043-1051. [doi: 10.1001/jama.294.9.1043] [Medline: 16145024]
13. Maloney S, Storr M, Morgan P, Ilic D. The effect of student self-video of performance on clinical skill competency: A randomised controlled trial. Adv Health Sci Educ Theory Pract 2013 Mar;18(1):81-89. [doi: 10.1007/s10459-012-9356-1] [Medline: 22354337]
14. Maloney S. Investigating the efficacy of practical skill teaching: A pilot-study comparing three educational methods. Adv Health Sci Educ Theory Pract 2012:71-80. [doi: 10.1007/s10459-012-9355-2]
15. Moorhead S, 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]
16. Maloney S, Moss A, Ilic D. Social media in health professional education: A student perspective on user levels and prospective applications. Adv Health Sci Educ Theory Pract 2014 Dec;19(5):687-697. [doi: 10.1007/s10459-014-9495-7] [Medline: 24566977]
17. Maloney S, Chamberlain M, Morrison S, Kotsanas G, Keating JL, Ilic D. Health professional learner attitudes and use of digital learning resources. J Med Internet Res 2013;15(1):e7 [FREE Full text] [doi: 10.2196/jmir.2094] [Medline: 23324800]
18. Rosenthal S, Howard B, Schlussel YR, Herrigel D, Smolarz BG, Gable B, et al. Humanism at heart: Preserving empathy in third-year medical students. Acad Med 2011 Mar;86(3):350-358. [doi: 10.1097/ACM.0b013e318209897f] [Medline: 21248596]
19. Rio E, Moseley L, Purdam C, Samiric T, Kidgell D, Pearce AJ, et al. The pain of tendinopathy: Physiological or pathophysiological? Sports Med 2014 Jan;44(1):9-23. [doi: 10.1007/s40279-013-0096-z] [Medline: 24027089]
20. Hutchinson L. Evaluating and researching the effectiveness of educational interventions. BMJ 1999 May 8;318(7193):1267-1269 [FREE Full text] [Medline: 10231262]
21. Wang AT, Sandhu NP, Wittich CM, Mandrekar JN, Beckman TJ. Using social media to improve continuing medical education: A survey of course participants. Mayo Clin Proc 2012 Dec;87(12):1162-1170 [FREE Full text] [doi: 10.1016/j.mayocp.2012.07.024] [Medline: 23141117]
22. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med 2012 Oct 4;367(14):1355-1360 [FREE Full text] [doi: 10.1056/NEJMsr1203730] [Medline: 23034025]
23. GenStat Release 17 Reference Manual, Part 2 Directives. Hemel Hempstead, UK: VSN International; 2014.
24. Lumley T, Diehr P, Emerson S, Chen L. The importance of the normality assumption in large public health data sets. Annu Rev Public Health 2002;23:151-169. [doi: 10.1146/annurev.publhealth.23.100901.140546] [Medline: 11910059]
25. Ilic D, Nordin RB, Glasziou P, Tilson JK, Villanueva E. A randomised controlled trial of a blended learning education intervention for teaching evidence-based medicine. BMC Med Educ 2015;15:39 [FREE Full text] [doi: 10.1186/s12909-015-0321-6] [Medline: 25884717]
26. Carvas M, Imamura M, Hsing W, Dewey-Platt L, Fregni F. An innovative method of global clinical research training using collaborative learning with Web 2.0 tools. Med Teach 2010;32(3):270. [doi: 10.3109/01421591003719525] [Medline: 20218845]
27. Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-Based Medicine: How to Practice and Teach It. 4th edition. Edinburgh, UK: Churchill Livingstone; 2011.

Abbreviations

REML: restricted maximum likelihood
SMUPI: Social Media Use and Perception Instrument

Edited by G Eysenbach; submitted 28.05.15; peer-reviewed by E Dafli, M Connelly; comments to author 22.06.15; revised version received 17.09.15; accepted 30.09.15; published 26.10.15.

Please cite as:
Maloney S, Tunnecliff J, Morgan P, Gaida JE, Clearihan L, Sadasivan S, Davies D, Ganesh S, Mohanty P, Weiner J, Reynolds J, Ilic D
Translating Evidence Into Practice via Social Media: A Mixed-Methods Study
J Med Internet Res 2015;17(10):e242
URL: http://www.jmir.org/2015/10/e242/
doi:10.2196/jmir.4763
PMID:26503129

Original Paper

Medical Content Searching, Retrieving, and Sharing Over the Internet: Lessons Learned From the mEducator Through a Scenario-Based Evaluation

Athos Antoniades[1,2*], PhD; Iolie Nicolaidou[1,3*], PhD; Dimitris Spachos[4], BSc, MSc; Jarkko Mylläri[5], MEd; Daniela Giordano[6], PhD; Eleni Dafli[4], MSc, MD, PhD; Evangelia Mitsopoulou[7], BSc, MSc; Christos N Schizas[1], PhD; Constantinos Pattichis[1], PhD; Maria Nikolaidou[4], MSc; Panagiotis Bamidis[4], PhD

[1]Department of Computer Science, University of Cyprus, Nicosia, Cyprus

[2]STREMBLE Ventures Ltd., Limassol, Cyprus

[3]Department of Communication and Internet Studies, Cyprus University of Technology, Limassol, Cyprus

[4]Lab of Medical Physics, Medical School, Faculty of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece

[5]Department of Teacher Education, University of Helsinki, Helsinki, Finland

[6]Dipartimento di Ingegneria Elettrica, Elettronica e Informatica, Università di Catania, Catania, Italy

[7]e-Learning Unit (ELU), Division of Population Health Sciences & Education, St George's, University of London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Panagiotis Bamidis, PhD
Lab of Medical Physics
Medical School, Faculty of Health Sciences
Aristotle University of Thessaloniki
PO Box 376
Thessaloniki, 54124
Greece
Phone: 30 2310 999310
Fax: 30 2310 999702
Email: bamidis@med.auth.gr

Abstract

Background: The mEducator Best Practice Network (BPN) implemented and extended standards and reference models in e-learning to develop innovative frameworks as well as solutions that enable specialized state-of-the-art medical educational content to be discovered, retrieved, shared, and re-purposed across European Institutions, targeting medical students, doctors, educators and health care professionals. Scenario-based evaluation for usability testing, complemented with data from online questionnaires and field notes of users' performance, was designed and utilized for the evaluation of these solutions.

Objective: The objective of this work is twofold: (1) to describe one instantiation of the mEducator BPN solutions (mEducator3.0 - "MEdical Education LINnked Arena" MELINA+) with a focus on the metadata schema used, as well as on other aspects of the system that pertain to usability and acceptance, and (2) to present evaluation results on the suitability of the proposed metadata schema for searching, retrieving, and sharing of medical content and with respect to the overall usability and acceptance of the system from the target users.

Methods: A comprehensive evaluation methodology framework was developed and applied to four case studies, which were conducted in four different countries (ie, Greece, Cyprus, Bulgaria and Romania), with a total of 126 participants. In these case studies, scenarios referring to creating, sharing, and retrieving medical educational content using mEducator3.0 were used. The data were collected through two online questionnaires, consisting of 36 closed-ended questions and two open-ended questions that referred to mEducator 3.0 and through the use of field notes during scenario-based evaluations.

Results: The main findings of the study showed that even though the informational needs of the mEducator target groups were addressed to a satisfactory extent and the metadata schema supported content creation, sharing, and retrieval from an end-user perspective, users faced difficulties in achieving a shared understanding of the meaning of some metadata fields and in correctly managing the intellectual property rights of repurposed content.

Conclusions: The results of this evaluation impact researchers, medical professionals, and designers interested in using similar systems for educational content sharing in medical and other domains. Recommendations on how to improve the search, retrieval, identification, and obtaining of medical resources are provided, by addressing issues of content description metadata, content description procedures, and intellectual property rights for re-purposed content.

(J Med Internet Res 2015;17(10):e229) doi:10.2196/jmir.3650

KEYWORDS

searching and sharing of medical educational content; repurposing; metadata; evaluation

Introduction

The mEducator Best Practice Network and Solutions Proposed

Although there is an abundance of medical educational content available in individual academic institutions, it is not widely available or easy to discover and retrieve due to a lack of standardized content-sharing mechanisms. Medical education institutions often use a variety of Web-based Learning Content Management Systems (LCMSs) to support the teaching and learning process. They also use some of the available educational standards to describe the educational content in the LCMS. This would presumably allow for better managing of the content through the Web and increase its interoperability in different LCMSs and platforms, but instead these learning materials often remain confined within the individual institutions. Sharing educational material has been the focus of recent developments and practice. Web 2.0 has highlighted the importance of sharing, as well as social collaboration and participation, social networking, and crowd intelligence in the domain of health education, where medical doctors, medical students, practitioners, and others, can benefit from open access to information and sharing of ideas, questions, and opinions [1].

Despite numerous efforts and emphasis in the area of educational content development and its sharing through social media (eg, YouTube), there has been no prominent, clear, and standards-based solution for the seamless sharing of educational content in medicine, where seamless implies immediate awareness of any new educational content, accuracy of its classification regarding the topics and skills being addressed, and visibility of any existing adaptations of the content to suit different educational context (repurposing). For example, recent research has examined the use of a standardized medical thesaurus (SNOMED CT) in YouTube health video tags from preselected YouTube medical education channels and found that the average percentage of YouTube tags expressed using SNOMED CT terms was about 22% [2]. To fill this gap, the mEducator Best Practice Network (BPN) implemented standards and reference models in e-learning to develop innovative solutions that enable specialized state-of-the-art medical educational content to be discovered, retrieved, shared, and re-used across European Institutions, targeting medical students, doctors, educators, and health care professionals [3]. For this aim, mEducator also had to elaborate on pedagogical, technical, standardization, cultural, social, and legal issues.

mEducator has tackled the challenge of seamless sharing of educational content by fusing the social Web and the semantic Web concepts. Learning management systems and open educational repositories were united so that educators and learners could organize, repurpose (defined as convert for use in another format or educational context), re-use, and share medical educational resources. Different platforms have been created to enable the organization, repurposing, re-use, and sharing of medical educational resources. The mEducator ontology has played a pivotal role in this endeavor, as it has been designed to provide the various mEducator instances with a metadata schema that has well-defined semantics [1]. These instances, all using the same metadata schema, can be classified based on the specific underlying technologies: two solution frameworks that were developed for multitype content sharing and repurposing.

First Solution: mEducator2.0, Based on Web 2.0 Technologies (Specifically Mashups)

In mEducator2.0, a brokerage mechanism was created based on mashups and other Web 2.0 technologies, which allows medical educational content to be shared across LCMSs, thereby creating a loosely coupled network of LCMSs. A mashup is a Web application that uses content from more than one source to create a single new service displayed in a single graphical interface. Users can access mEducator educational material through the mEducator2.0 portal and their own systems using mashup technologies. mEducator users across multiple institutions may use the mashups for uploading, creating, and editing content metadata, as well as for the search and retrieval of content. Alternatively, for users without access to a specific LCMS, an independent platform has been created, which applies Web 2.0 techniques and facilitates user collaboration, allowing, at the same time, the creation of social networks for medical education and knowledge exchange [1].

Second Solution: mEducator3.0, Based on Semantic Web Technologies (Specifically Linked Data)

The mEducator3.0 solution is designed around a federated architecture based on a service-oriented application framework and the use of semantic technologies. The framework is fundamentally based on the Semantic Web Services-oriented e-learning architecture [4,5] and the emerging Linked Data and Linked Services paradigms [6]. This solution fundamentally exploits semantic representations of data and services to provide interoperability between e-learning repositories spread across the Web. In particular, descriptions of the content are based on machine-understandable metadata and vocabularies that also follow linked data principles [7,8].

There are four instantiations of the mEducator3.0 solution, which differ in the type of content management technologies

(CMT) they are integrated with: (1) mEducator3.0-MILES+, based on semantic extensions of Moodle [9,10], (2) mEducator3.0-MELINA+, based on the Drupal content management system (CMS), (3) mEducator3.0-Linked Labyrinth+ [11], based on the Open Labyrinth platform for Virtual patients, and (4) mEducator3.0-Metamorphosis+, based on the social network, Elgg [12,13].

This paper focuses on the evaluation of MELINA+, with respect to the underlying metadata schema and the overall usability and user acceptance of the system. We chose to report on this solution because it was the most widely tested, in a number of different settings, in four different countries (ie, Greece, Cyprus, Bulgaria, and Romania), with a sufficient number of users. This study is original in implementing scenario-based assessment for evaluating content-sharing technologies in tandem with end-users, contents, and their interactions. The paper is structured as follows: the next section synthesizes the findings of a literature review on evaluating metadata in e-learning systems, which leads to the section where we define the term "metadata schema" and illustrate its importance in the context of mEducator. We then provide the rationale for using scenario-based evaluation to address the target users' needs and conclude with the research goal of the study. The methodology section describes how scenario-based evaluation was conducted and provides information on the participants of the study and the instruments that were used. The results section presents the results of the evaluation of the metadata schema and the overall usability of the system, as well as how the latter was perceived and accepted by users. The discussion section provides recommendations and lessons learned from the evaluation and draws implications for the design of content sharing and repurposing systems for medical education.

What We Know From Previous Research

Evaluating Metadata in E-Learning Systems: Research Background and Rationale

The evaluation framework of mEducator has drawn from the literature on the implementation and evaluation of metadata in other e-learning systems [14]. The term metadata is used to describe data that provide additional information about a certain item's content, that is, data about data. A metadata schema is composed of a set of terms, a set of structural definitions of metadata instances, and a binding schema for implementation [15,16]. In the e-learning context, metadata, by providing descriptive information about resources and learning objects, facilitate retrieval and re-use in various instructional contexts and have been researched extensively [17]. One such example is the Adaptive Hypermedia Knowledge Management E-Learning Platform, which used metadata to satisfy requirements such as reusability and interoperability and provided tools to support teachers in the evaluation, import, and retrieval of high-quality educational resources. Rego et al [17] focused on the evaluation of the quality of learning objects and developed a tool that involved the use of an intelligent agent using data mining techniques for the analysis of the metadata contained in the learning object.

Researchers, who concentrated their efforts on designing a system where users could most effectively find the items they searched for, agreed on two crucial factors: (1) the effective use of metadata [18] and (2) a user-centered approach [19-21]. Morales-Salcedo et al [19] in their U-campus (Ubiquitous campus) project, provided users with means to access and control all available resources in a uniform fashion from a single vantage point. Gkatzidou et al [20] adopted a user-centered approach in their effort to create reusable, accessible, and adaptable learning objects, by focusing attention to user profiles. In their case, metadata were used to describe not only the learning object but also the learner's profile. Specifically, the description of the user's immediate needs and preferences was matched with a description of the components of a resource or service to provide an accessible relationship between the learner and the resource and enable the delivery of learning content that has been adapted to suit the needs of the individual user. However, they did not provide a way for evaluating their approach.

To reduce the subjectivity in the evaluation of the quality of metadata, some researchers [22] borrowed the evaluation framework proposed by others [23] and provided methodological guidelines that are useful for the evaluation of metadata in other types of systems. This framework summarized the quality of the metadata instance in seven measurable parameters: (1) completeness, (2) accuracy, (3) provenance, (4) conformance to expectation, (5) logical consistency and coherence, (6) timeliness, and (7) accessibility. As explained in [22]:

> *In a complete metadata record, the learning object is described using all the fields that are relevant to describe it. In an accurate metadata record, the data contained in the fields correspond to the object that is being described. The provenance parameter reflects the degree of trust that you have in the creator of the metadata record. Conformance to expectations measures how well the data contained in the record let you gain knowledge about the learning object without actually seeing the object. Logical consistency and coherence reflects two measures: The consistency measures if the values chosen for different fields in the record agree between them. Coherence measures if all the fields talk about the same object. Timeliness measures how up-to-date the metadata record is compared with changes in the object. Accessibility measures how well you are able to understand the content of the metadata record (p. 6).*

Adopting a user-centered approach and recognizing both the limitations of having experts create metadata and the limitations of automatically generating them, Zens and Baumgartner [21], in their effort to allow users to find resources that fit their needs, suggested social tagging of metadata in the context of education as a third way of creating metadata. The pedagogical potential of social bookmarking was also critically presented by Dias et al [24]. The system by Zens and Baumgartner [21] was called MELT (A Metadata Ecology for Learning and Teaching), and it represents a content enrichment project that bridged 17 public and private sector content partners with the goal of promoting the exchange of learning resources across Europe. MELT used

an existing brokerage system that supported federated searching across a network of linked content repositories. MELT pursued a multilayer metadata enrichment approach that included expert indexing, automatic metadata generation, and social tagging. Social tagging allowed users to add tags to given objects and thus reflected the view of multiple users. It resulted in many accumulative metadata records related to a given resource. Social tagging results in folksonomies. Folksonomies are user-generated taxonomies that facilitate the sharing of content within a social network of users and potentially promote an efficient discovery of learning resources that meet the user's needs. The potential of taxonomies for learning objects in the medical domain has been explored by mEducator in [25]. A similarity between MELT and mEducator is that mEducator also allows users to add open-ended metadata fields when adding content in the platform [25], when the existing metadata do not cover a specific aspect they consider important. For the evaluation of the effectiveness of MELT, seven success indicators were implemented, some of which are applicable in the context of mEducator, as well. Those success indicators are the following: (1) effectiveness and efficiency of the search and retrieval process, (2) utility of metadata enriched by experts for finding relevant content, (3) utility of folksonomies for finding relevant content, (4) effectiveness of automatically generated metadata regarding discovery of content, (5) user satisfaction, (6) use of the retrieved content, and (7) use of content across languages and across countries.

Evaluating Metadata in mEducator

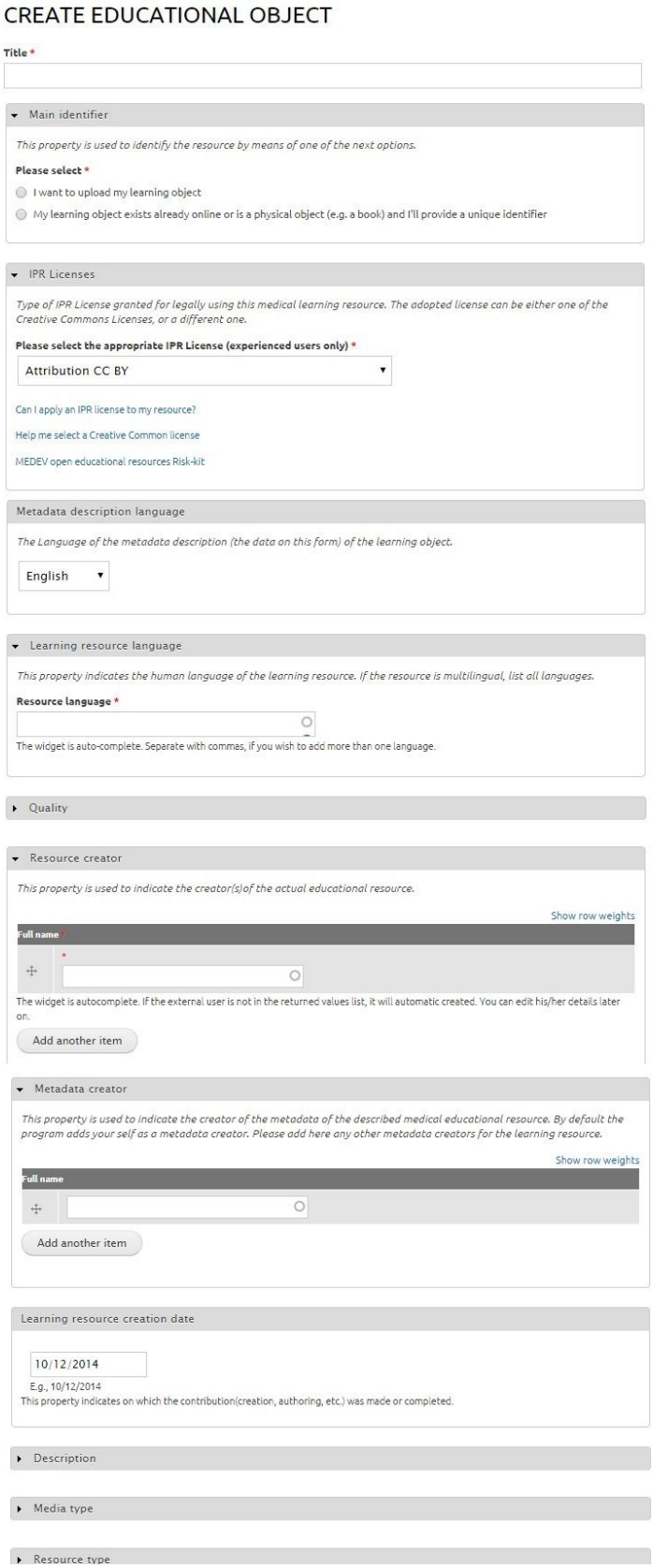
The term "metadata schema" is used in mEducator to describe medical educational resources of various types in a standardized, machine processable format in order to enable the medical educational resources to be shared, exchanged, searched, and retrieved across academic institutions. One key challenge was that the completion of metadata is not performed by professional indexers, as was the case in previous research [26], but is largely delegated to end-users and content providers. This is challenging because, in general, it is difficult for the end-users to achieve a shared understanding of the meaning of the metadata fields, as demonstrated in studies of digital libraries [27] and in some preliminary studies carried out within mEducator [28].

The approach followed by mEducator in the design of the metadata schema was based on an attempt to balance competing requirements. Specifically, the number of required fields, including ones deemed essential to the mEducator content-sharing model, that is, specification of the intellectual property rights (IPR) for the shared content and the description of the content modification carried out during repurposing extensions was minimized, while the opportunity for richer descriptions of the resource was still provided in the system. Controlled vocabularies/taxonomies were introduced to facilitate classification of the educational resource. A set of additional optional fields were proposed for a more comprehensive description of the educational/pedagogical aspects of the content. The metadata-filling interface in mEducator MELINA+ is shown in Figure 1. After providing a title, a user had to specify fields, such as IPR licenses, metadata description language, learning resource language, quality, resource creator, metadata creator, learning resource creation date, description, media type, resource type, and discipline (Figure 1).

The evaluation of the mEducator3.0 solution MELINA+ was based on an evaluation framework specifically developed for assessing the effectiveness of all the platforms that were implemented through the mEducator BPN and any future platforms that attempt to address the same issues. An a priori analysis of the goals and nature of the mEducator solutions pointed to five important pillars that need to be addressed in an evaluation model [29]: (1) IPR of content (refers to license types and mechanisms of content protection), (2) repurposing (refers to the tracking of content genealogy and content re-use activities), (3) accessibility of metadata schema, (4) content evaluation (system generated analyses, user review, peer review process), and (5) content sharing. At the same time, three dimensions are transversal and span those pillars: (1) human-computer interaction (HCI) (overall accessibility, usability, consistency of the proposed solutions), (2) technological issues (requirements for content providers, handling of content updates, system performance, and maintenance requirements), and (3) sustainability (incentives for content providers/ consumers/ reviewers and financial sustainability).

mEducator attempted to take the context of the target users' activities into account in the evaluation effort. Accordingly, the targeted user groups serve as the point of departure for designing and executing user testing. Evaluation is based on scenarios typical for each user group's context of using Web services (scenario-based evaluation). These activity scenarios also take into account the dominant or typical content types for each user group and context. The decision to use activity scenarios is supported by previous research that showed that good assessment scenarios are particularly revealing and valuable because they ask learners to make decisions by applying their understanding of the system [30].

Figure 1. The metadata filling interface in meducator MELINA+.



Scenario-Based Evaluation: Research Background and Rationale

Several research attempts have focused on developing a scenario-based evaluation model in diverse disciplines, such as the online assessment of students' problem solving skills [30], educational assessment [31], service-oriented architectures [32], and software engineering [33], to name a few. In medical education in particular, scenario-based assessment is being used extensively for the assessment of clinical skills [34] or for

training [35] or even for exploring design requirements for repurposing medical cases [36].

In the context of educational assessment, Mayotte [31] attempted to evaluate the effectiveness of an interactive scenario-based assessment system to address the limitations of traditional assessment methods, allowing students to troubleshoot complex scenarios, ask questions, and make diagnoses through an interactive Web interface. His research showed that scenario-based assessment placed greater emphasis on learners' problem solving, critical thinking, and reasoning skills. In the context of educational assessment within medical education, Nestel et al collected performance data [34] when third-year medical students worked through scenarios undertaking defined tasks for the assessment of their procedural skills in contextualized (specific objective-driven) tasks. In the latter research [34], it was found that scenario-based evaluation was valued for the opportunity to practice patient-centered care in a simulated setting that integrated technical, communication, and other professional skills. These researchers concluded that scenario-based assessment reflected real-world issues of patient-centered care [34]. In the context of scenario-based training in medical education, similar findings were reported by [35] who found that scenario-based virtual world team training of cardiopulmonary resuscitation (CPR) was feasible and showed promising results for the training of medical students in multiperson CPR. Moreover, the value of scenario-based think aloud protocols in the area of symptom interpretation, online diagnosis, and HCI was demonstrated by [37] in their work with adults aged over 50 years.

To the best of our knowledge, no research studies on scenario-based assessment for evaluating content-sharing technologies in tandem with end-users, contents, and their interactions exist in the literature [38].

Research Goal

The focus of the evaluation methodology was to define the extent to which the informational needs of the mEducator target users were covered and if the schema supported content searching, retrieval, and sharing from an end-user perspective. For example, the construct "metadata accuracy" was operationalized to address the users' notion of whether the metadata are understandable, whether and how they cover the informational needs of the mEducator target group, and how they function to retrieve relevant content. The term "accuracy" is therefore referred to in this study in its broader sense of general adequacy for the target users' needs, rather than in the more technical acceptance of the term, often used in the literature, to indicate whether in any digital collection the metadata are filled with accurate/precise information with respect to the digital object they refer to.

The research goals of our study were to assess (1) how the metadata schema developed by mEducator addresses the informational needs of end-users in the process of medical content searching, (2) how the metadata schema developed by mEducator supports the process of content sharing, (3) the overall usability of the schema as implemented in the mEducator3.0 MELINA+ system, and (4) the overall user acceptance of this particular mEducator system.

The evaluation relates to the medical domain in three main ways: (1) some fields of the metadata schema refer to the medical/health domain and thus are very specific, (2) the testing scenarios used in the evaluations are largely dominated by medical/health-related cases in which the aforementioned medical-domain–related metadata are involved, and (3) the whole evaluation context was medical, in the sense that participants were either medical students (undergraduate or post-graduate) or health professionals. So in this study, (1) medical-domain specific metadata were defined and used, (2) medical-domain–related scenarios were constructed and used, and (3) all participants were from the medical domain.

Methods

Scenario-Based Evaluation Context

A number of assessment methods, including scenario-based assessment for usability testing complemented with data from online questionnaires and field notes, were designed and used to evaluate the different instantiations of the mEducator solutions. Overall, the evaluation effort in mEducator addressed two main levels: (1) the system perspective, that is, technical evaluation, based on functional requirements, and (2) the usage perspective, that is, scenario-based evaluation. Both were carried out with a focus on quality of service and user experience. The scenario-based evaluation across user groups was decided with a special emphasis on the user testing stage. A framework for integrated scenario-based and quality of service evaluation was developed (Figure 2). The overall evaluation framework for the mEducator system was divided into technical evaluation, which was based on functional requirements, and scenario-based evaluation, which is the focus of this study.

As Figure 2 shows, "scenario-based evaluation" was developed for the individual "users" and "institutions" but also addressed technical issues, always taking into consideration the target users' ("teachers", "students", and "doctors") roles as "contributors" or "consumers" of medical educational material in the mEducator system (Figure 2). Case studies examining issues of pedagogy and adoptability in individual institutions, as well as technical issues of LCMS integration and repository characteristics, which are shown in Figure 2, are beyond the scope of this paper. It is important to note that different data sources and methods were used for the different levels of the evaluation. For example, in the scope of the overall study, interviews, observational case studies, screenshot capturing, and automated tracking of activity were implemented, primarily in the initial stage of the project to inform and evaluate the first prototypes of the mEducator system.

Scenario-based evaluation allowed the examination of key areas of the metadata schema, such as the following: (1) capability to embrace all types of medical learning resources, (2) capability to account for content repurposing, and (3) handling copyright licenses, especially of the parent resource during repurposing.

Evaluation workshops, referred to as case studies in this paper, were organized and run at various mEducator partners' sites. This paper focuses on the analysis of four such case studies, through which the mEducator evaluation framework was

applied. These case studies were conducted between August 31, 2011, and April 7, 2012, and they focused on one mEducator instantiation, MELINA+, which provided options to users to explore a large collection of educational objects, create their own educational object including a full metadata description, and collaborate with other members of the community.

More specifically, MELINA+, which stands for MEdical Education LINnked Arena, is a CMS for medical educational resources, based on Drupal 7, an open source content management system. MELINA+ has been developed to allow resources to be created, uploaded, described, shared, and searched over the semantic Web in different ways. It exploits SPARQL queries using the Drupal SPARQL endpoint functionality. Multiple endpoints (internal and external) could be added in SPARQL registry and queried. It supports the creation and description of learning resources, user registration/authentication, advanced search capabilities, a commenting/rating/bookmarking system, blogs, and posts. Its advanced features include core RDF support, embedded SPARQL endpoint, DBpedia spotlight annotation, social learning collaboration, quality process control for learning resources, and single sign-on via WEBID module. Figure 3 shows the interface of MELINA+, specifically users' options to "explore", "create", and "collaborate".

Evaluation sessions started with a demonstration of the mEducator system conducted by a facilitator. It is important to clarify that scenario-based evaluation involves a facilitator that guides participants, especially when the latter are not familiar with a particular system, and as such it is a method typically applied in structured training and evaluation sessions. In the four case studies described in this paper, a facilitator and 2 research assistants were present to guide participants and also to take field notes based on their observations concerning the points where the target users faced difficulties or expressed doubts. Research assistants were instructed to take note of explicit requests for clarification when these were made by participants and also to take notes of the context in which the users seemed to be stuck or puzzled while performing each scenario.

In the context of scenario-based evaluation, participants typically created a learning resource, for example a presentation outlining the aims and results of a medical paper, complemented with educational information, such as educational objectives, target audience, and expected learning outcomes.

Participants had access to a number of different types of learning content in mEducator, in different languages, including text, multimedia slide presentations, images, videos, interactive learning environments, wikis, sketch graphical annotations, virtual patients, websites, animations, audio files, 3D models, and eBooks, all in the medical domain. Each case study included three scenarios, as described below.

Figure 2. Overview of the evaluation framework developed for the mEducator system, dividing evaluation into technical evaluation and scenario-based evaluation.

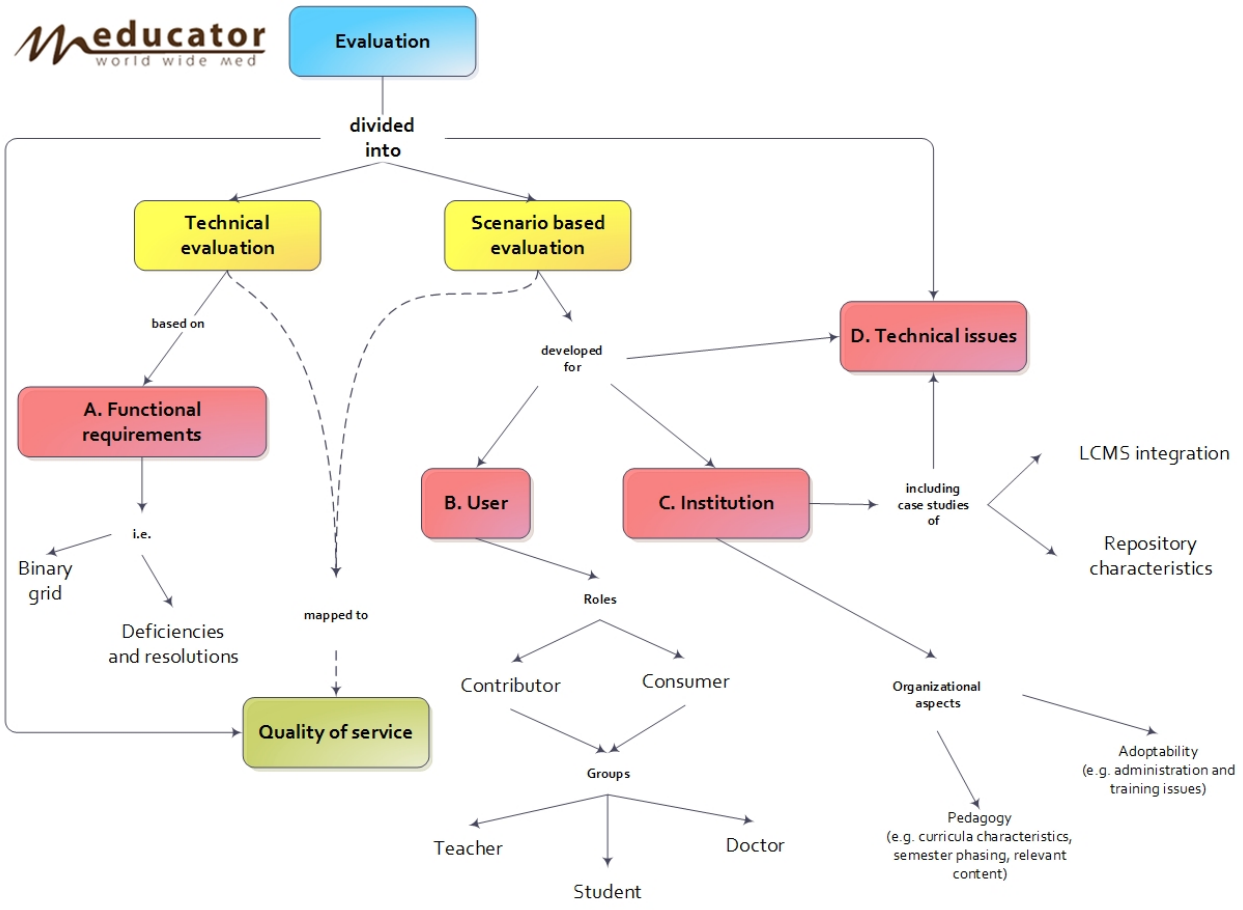
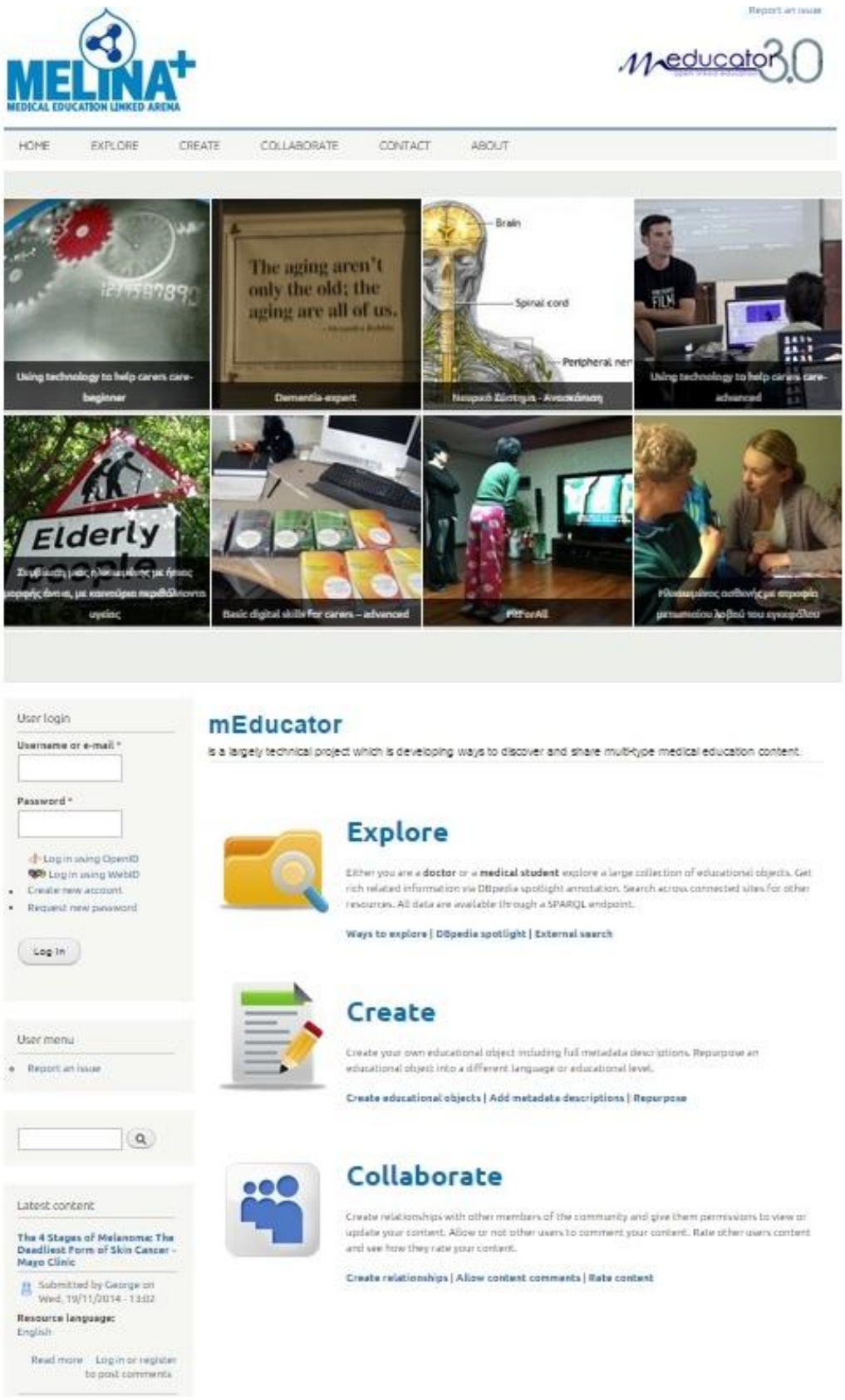


Figure 3. Screenshot of the interface of mEducator3.0 MELINA+ instantiation.



Scenario 1: Creating an Educational Medical Resource and Its Metadata

The first scenario referred to the *creation* of an educational resource in MELINA+ (Figure 3, “Create”). Participants created a new account, logged into the system, uploaded a learning resource, entered and saved the appropriate metadata for the resource, and visualized the entered metadata in the system. Before they were in a position to evaluate MELINA+, they had to create an educational resource in mEducator3.0 MELINA+

and familiarize themselves with the medical educational resource creation processes. The role of the facilitator in Scenario 1 was to (1) explain the basic functionalities involved in medical educational resource creation, (2) explain the notion of metadata, especially those describing the resource in medical terms, and (3) help users appreciate the notion of the IPR metadata used for describing medical educational resources when it comes to sharing and repurposing medical content, even for educational uses.

Scenario 2: Searching for a Medical Educational Resource

The second scenario referred to the process of *searching* for a resource in MELINA+ (Figure 3, "Explore"). Before they were in a position to evaluate MELINA+, participants were expected to specify the search attributes, perform the search, and analyze the obtained results. The aim of Scenario 2 was to familiarize users with the searching of resources in mEducator3.0 MELINA+. The role of the facilitator in Scenario 2 was to (1) demonstrate to medical users how to search for educational resources in comparison to simple search engines, (2) explain the role of search attributes, and (3) help medical users appreciate notions of metadata used for describing the medical resource (and its subsequent sharing with other medical users).

Scenario 3: Repurposing and Specifying Intellectual Property Rights Attributes for Sharing Medical Resources

The third scenario referred to *repurposing* a resource and treating the IPR aspects of it (Figure 3, "Collaborate"). It is important to note that, by the third scenario, participants have already encountered the notion of IPR, as clearing IPR is compulsory in creating and sharing a resource. Before they were in a position to evaluate MELINA+, participants were expected to understand the notion of IPR and creative commons, specify the IPR metadata and attributes for their own resources, perform some kind of repurposing, correctly fill in repurposed resources, and analyze the obtained results. The role of the facilitator in Scenario 3 was to (1) help medical users familiarize themselves with the aspects of medical resource repurposing in mEducator3.0 MELINA+, (2) help medical users familiarize themselves with the notion of utilizing/exploiting metadata to link resources in an hierarchic way suitable for sharing in the medical education domain, and (3) explain to medical users the basic notion of repurposing in practice and its importance in medical education.

It is important to note that too few of the participants were able to complete the third scenario that referred to repurposing with regard to sharing within the allocated time. Therefore, this study focused primarily on creating, searching, and retrieving medical educational resources. Another important note is that the concept of repurposing is mostly of interest to professors and medical professionals, rather than to medical students, who constituted the majority of the participants in this study.

At the end of their interaction with MELINA+, participants completed two evaluation instruments, which are described in detail in the next section. Participants also completed two open-ended questions, which focused on the perceived system's strengths and weaknesses.

Participants

Overall, there were 126 respondents who evaluated the mEducator3.0 instantiation MELINA+. The majority of these respondents (77.8%) were undergraduate students, followed by 8.7% of graduate students, 6.3% of postgraduates, 4% of medical professionals, and 3.2% of professors.

Instruments

Two main instruments were used as data sources for the mEducator evaluation framework: (1) the "System Usability Scale (SUS) questionnaire" and (2) the "Questionnaire on the metadata and search process". The first instrument was primarily used to assess the overall usability of the system in absolute terms. The second instrument was primarily used to document (1) the extent to which the metadata schema developed by mEducator addresses the end-users' informational needs in the process of medical content searching and (2) the extent to which the metadata schema developed by mEducator supports the process of content sharing. Both instruments were completed online in the form of an online anonymous survey.

Moreover, another source of data referred to the facilitator's and research assistants' field notes during the MELINA+ evaluation sessions, which documented cases where users expressed doubts or faced difficulties during the execution of the scenarios. These field notes were primarily used as a complementary source of data for triangulation purposes and also to help derive recommendations for the improvement of the system, based on the users' input. The analysis of the first two data sources, the two instruments, includes (1) quantitative results from the completion of online surveys, and (2) qualitative results of open-ended questions in these surveys.

The System Usability Scale Questionnaire

The System Usability Scale (SUS) questionnaire (see Multimedia Appendix 1 for full instrument) provides a well-known, widely used, and standardized quantifiable measure for the usability of systems, which is provided in the form of an absolute scale that allows comparisons between systems. Selecting a standardized method that would allow the evaluation and comparison of the different instantiations of mEducator was a critical factor that had to be considered during the mEducator project.

The SUS questionnaire consists of 10 questions alternating in positive and negative phrasing to remove any bias arising from the phrasing of questions [39]. Each question is answered on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). An advantage of the SUS questionnaire is that it produces a standardized score ranging from 0-100 that can be used to compare the usability of systems directly [40].

With regard to how the results of the SUS are analyzed, this instrument yields a single number representing a composite measure of the overall usability of the system being studied. Scores for individual items are not meaningful on their own. The procedure followed for scoring the SUS questionnaire and the results was the following: for odd questions, 1 was subtracted

from the user response, and for even-numbered questions, the user response was subtracted from 5. This scales all values from 0-4 (with 4 being the most positive response). The converted responses were added up for each user and the total was multiplied by 2.5. This converted the range of possible values from 0-100 instead of 0-40.

Questionnaire on the Metadata and Search Process

The second instrument consisted of 16 questions that focused on the metadata and search process evaluation (see Multimedia Appendix 2 for full instrument). These questions examined metadata (Q2-4), the usefulness and relevance of retrieved content (Q1, Q5-8, Q15-16), latency and difficulty level of searches (Q9, Q12-14), IPR (Q11), and open sources (Q10).

Each question was answered on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). The questionnaire, which was administered in an online format, was followed by two open-ended questions: (1) What are the positive and negative aspects of the system for searching and finding content?, and (2) What advantages and disadvantages do you perceive of a portal with international content?

Results

This section presents the results of the evaluation of the metadata schema and the overall usability of the system, as well as how this was perceived and accepted by users. The section starts with the results of the administration of the SUS and of the questionnaire on the metadata and search process. Following, a subset of the questions from the questionnaire is mapped into the International Federation of Library Associations and Institutions (IFLA) Framework [41], for an aggregated analysis. We then present the results of the analysis of the two open-ended questions.

Administration of the System Usability Scale

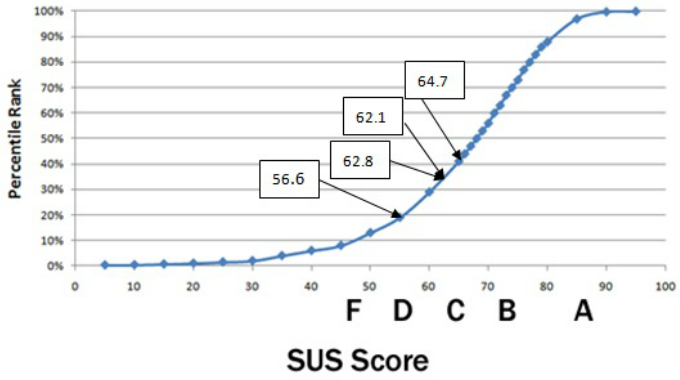
Table 1 presents the results of the evaluation using the SUS instrument from four different case studies that took place in four different countries (ie, Greece, Cyprus, Bulgaria, and Romania) between August 31, 2011, and April 7, 2012.

Table 1. SUS scores for the mEducator3.0 MELINA+ instantiation in four case studies.

Evaluation place	Evaluation as part of:	Group size/Profile	% SUS score for mEducator3.0/MELINA+
Plovdiv, Bulgaria	The e-Education and e-Science International Conference in Plovdiv October 5/6, 2011	25 medical professionals, students	56.6
Bucharest, Romania	An organized mEducator project event	15 medical students	62.8
Nicosia, Cyprus	An eHealth graduate class on November 11, 2011	35 postgraduate students taking an eHealth course	62.1
Thessaloniki, Greece	The Medical Education Informatics International Conference and Spring School on “Medical Education Content Sharing Technologies” on April 6/7, 2012	51 undergraduate medical students	64.7

Figure 4 shows how the SUS scores from the four case studies of the evaluation of MELINA+ are mapped on the SUS curve. Figure 4 shows that there is a progressive change of the overall system category from category D to C, over time, in the SUS score of the four case studies, starting from 56.6 in the first evaluation of MELINA+ in Bulgaria (category D) and reaching 64.7 in the fourth evaluation of mEducator MELINA+ in Greece (category C).

Figure 4. SUS scores of the four case studies of evaluation of mEducator 3.0 MELINA+ on the SUS curve.



Administration of the Questionnaire on the Metadata and Search Process

Results of this questionnaire constitute useful input for the analysis of the metadata accuracy to support the mEducator3.0 solution. It is important to note that the respondents' level of expertise varied as it ranged from undergraduate students to medical professionals and professors (see Participants section). However, the results of the different groups (eg, students as opposed to professors or medical professionals) did not differ substantially to justify the presentation of results per group of users. Therefore, results are presented for all participants in Table 2.

Table 2. Results for the evaluation of the metadata and search process (N=126).

Category examined	Question	N	Strongly disagree, %	Disagree, %	Neutral, %	Agree, %	Strongly agree, %
1. Metadata	Q2. The metadata is not understandable	126	7.9	53.2	25.4	8.7	4.8
	Q3. The amount of presented metadata is excessive	124	1.6	39.5	38.7	16.9	3.2
	Q4. The amount of presented metadata is insufficient	124	5.6	48.4	33.9	12.1	0
2. Retrieved content usefulness and relevance	Q1. The presented metadata helps me in revising the search or annotation terms	125	0	7.2	24	61.6	7.2
	Q5. I found useful content as outcome of my searches	120	2.5	1.7	17.5	55.8	22.5
	Q6. The amount of retrieved relevant content was adequate to my information needs	115	0	7.0	32.2	52.2	8.7
	Q7. The information immediately presented helps me assess the relevance of the resource	118	0.8	9.3	32.2	51.7	5.9
	Q8. I need to inspect the learning resource to assess its relevance	116	0	12.1	33.6	47.4	6.9
	Q15. I found interesting content outside the scope of my specific search	114	2.6	8.8	28.1	46.5	14
	Q16. I would recommend the system to my colleagues	123	1.6	4.1	14.6	49.6	30.1
3. Latency and difficulty level of searches	Q9. The search results were obtained quickly	118	0.8	0.8	14.4	63.6	20.3
	Q12. The advanced search form is easy to understand	115	0.9	7.8	24.3	57.4	9.6
	Q13. It is distracting to have international content listed in the results	118	4.2	36.4	42.4	16.1	0.8
	Q14. It was easy to inspect/download the (retrieved) learning resource	123	2.4	8.9	30.1	46.3	12.2
4. Assessing open sources	Q10. I could easily assess if the resource is open to use	123	1.6	5.7	34.1	53.7	4.9
5. IPR	Q11. It was difficult to understand the IPR of the resources	123	0	34.1	41.5	20.3	4.1

For the part of the questionnaire examining *metadata* (three questions), data analysis showed that in general more than half of participants agreed or strongly agreed that the metadata were understandable (17/126, 61.1%) and that the amount of metadata was sufficient (67/124, 54%).

With regard to the second category, *the usefulness and relevance of retrieved content* (seven questions), more than half of participants agreed or strongly agreed that the content they retrieved as an outcome of their searches was useful (94/120, 78.3%), adequate for their information needs (70/115, 60.9%), relevant (68/118, 57.6%), interesting (69/114, 60.5%), and also helped them in revising their searches (86/125, 68.8%). It is also important to note that a greater percentage, almost 79.7% (98/123) of participants, would recommend the system to their colleagues.

The third category referred to the *latency and difficulty level of searches* (four questions). The majority of participants agreed or strongly agreed that the search results were obtained quickly (99/118, 83.9%), the advanced search form was easy to understand (77/115, 67%), and that it was easy to download the retrieved learning resource (72/123, 58.5%). Last, more than half of participants agreed or strongly agreed that they could easily assess whether the learning resource was *open to use* (72/123, 58.5%). It must be noted that this aspect of the system was explicitly included in the questions because mEducator needed an instrument capable of registering any difference in performance of the various instantiations, given the different technologies (mashups and semantic searches) that were underlying the two classes of mEducator solutions.

The participants' response to a small number of questions was ambiguous and therefore difficult to interpret. What was

problematic in the findings was that in some questions (Q3, Q4, Q7, Q8, Q10, Q11, Q13), typically one in three participants was neutral and did not express an opinion. For example, with regard to whether metadata was excessive (Question 3), more than one third of participants (48/124, 38.7%) were neutral, while one fifth of them (25/124, 20.1%) agreed with the statement. It is possible that in this case, participants were uncertain of what the concept of "excessive metadata" meant.

One important issue that needs to be understandable by the user is the IPR license that has been assigned to the resource. The results of Q11 ("It was difficult to understand the IPR of the resources"), although overall positive, also provide an indication of confusion among users with respect to the term IPR. The majority of participants either disagreed (42/123, 34.1%) or were neutral (51/123, 41.5%) with respect to this question. Neutral responses can be interpreted in a number of ways. They might mean that users are uncertain of their answer. If the results are analyzed from this point of view, then the implication is that the presentation of IPR metadata needs to be improved. Regarding the evaluation of the IPR schema of the repurposed content, it can be analyzed from two different perspectives. The first one refers to the case where the to-be-repurposed resource is already published in the system, and the second one refers to the case where the to-be-repurposed resource is an external one. According to the first one, the mEducator schema successfully takes care of the licensing of the to-be-repurposed resource because every resource that is already in the system has a license. However, regarding the second case, the IPR license of the external resource is currently implemented in mEducator platforms but not in the schema. More specifically, there is no IPR-related field in the schema about the repurposed resource when this is an external one to the system. This latter situation is handled by the platform by requiring the user to tick a box, thus giving assurance that the user is authorized to make use of the learning resource. However, this information is not saved in the metadata.

The next section presents a more aggregated analysis, which is based on groups of questions mapped into a general reference model for bibliographic metadata.

Mapping the Online Survey into the International Federation of Library Associations and Institutions Framework: Aggregated Analysis

For convenience of interpretation, a subset of the questions from the online metadata accuracy survey was mapped into the IFLA Framework, namely the Functional Requirement for Bibliographic Record (FRBR) Framework [41]. The components of IFLA are (1) find, (2) identify, (3) select, and (4) obtain. Specific usability measures apply to these four generic tasks. Precision and helpfulness are important usability measures for the task of finding information. Ease of understanding, information adequacy, error rate, and information accuracy are important for both identifying and selecting information. Physical item accessibility is important for obtaining information. In this section, the responses to this subset of questions are analyzed, following the FRBR framework structure.

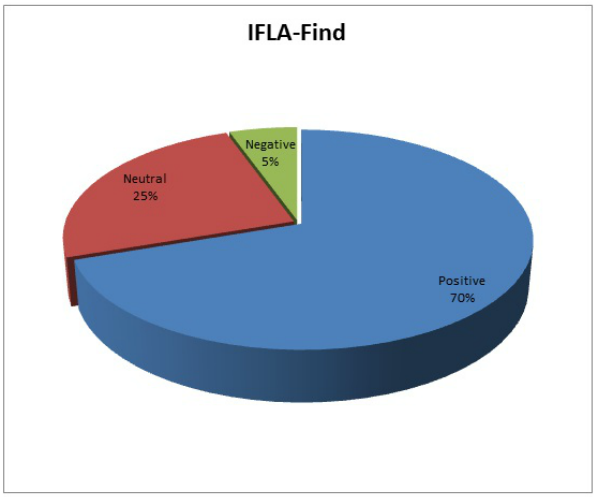
Find: Can a User Enter a Search and Retrieve Records Relevant to the Search?

By embedding the online metadata questionnaire into the IFLA Framework, the questions that were mapped under this stage of the IFLA framework are 5 and 6. Question 5 is "I found useful content as an outcome of my searches" and Question 6 is "The amount of retrieved relevant content was adequate to my information needs".

The results depicted in Figure 5 are based on a sample of 120 questionnaires and show whether a user can search and retrieve records relevant to the search. At first glance, it can easily be seen that 70.0% (84/120) of evaluators were satisfied regarding the retrieval of resources based on their needs. Not insignificant is the sum of neutral and negative results, which is 30.0% (25% +5%). Thus, approximately one third of evaluators (36/120) have either a negative or a neutral opinion on this.

Question 9, "The search results were obtained quickly", is also related to this IFLA stage because it investigates the speed with which results are retrieved. The average value of Question 9 was 4.01 and shows a predominantly positive attitude toward speed of retrieval, suggesting that there were no unacceptable delays during the retrieval of resources.

Figure 5. Users' responses mapped to the IFLA-Find requirement.



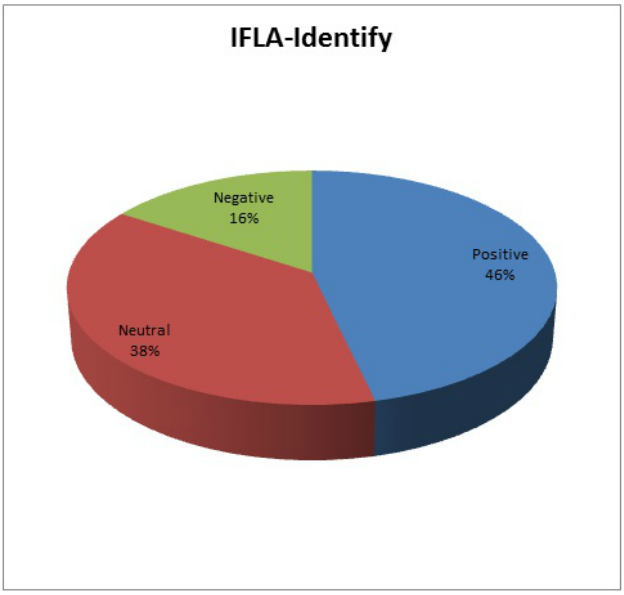
Identify: Once the User Retrieves a Record, Can They Successfully Interpret the Information in the Record to Know Whether the Source Information Will Be Relevant to their Needs?

The results of this section are informative, as they show to what extent the metadata are presented in a user-friendly way to the end-users and if they make sense to them, so that they are able to assess the suitability of metadata to their needs. The questions of the online metadata survey that were analyzed at this stage are Question 1: “The presented metadata helps me in revising the search or annotation terms”, Question 2: “The metadata is not understandable”, Question 4: “The amount of presented metadata is insufficient”, Question 7: “The information immediately presented helps me assess the relevance of the resource”, Question 10: “I could easily assess if the resource is open to use”, and Question 11: “It was difficult to understand the IPR of the resources”.

Figure 6 shows that 45.8% (55/120) of participants agree that they can interpret the retrieved information and can understand if it meets their needs or not. The remaining 54.2% (65/120) express a neutral to negative attitude. These results are not satisfactory and indicate that the meaning of metadata is not sufficiently clear to the evaluators.

Independent from the analysis of the questions above (1, 2, 4, 7, 10, 11), another question from the online metadata accuracy survey that is related to the current stage of IFLA is Question 3. It is worth looking at the average value of Question 3 because this metric shows the length of the metadata, which indirectly affects the ability of the learner to interpret the metadata. More specifically, the average value of Question 3, “The amount of presented metadata is excessive”, is 2.8. That means that most people disagree that the amount of metadata is excessive. This is a positive remark because it encourages them to further investigate and look at the resource, so they can interpret it better.

Figure 6. Users' responses mapped to the IFLA-Identify requirement.

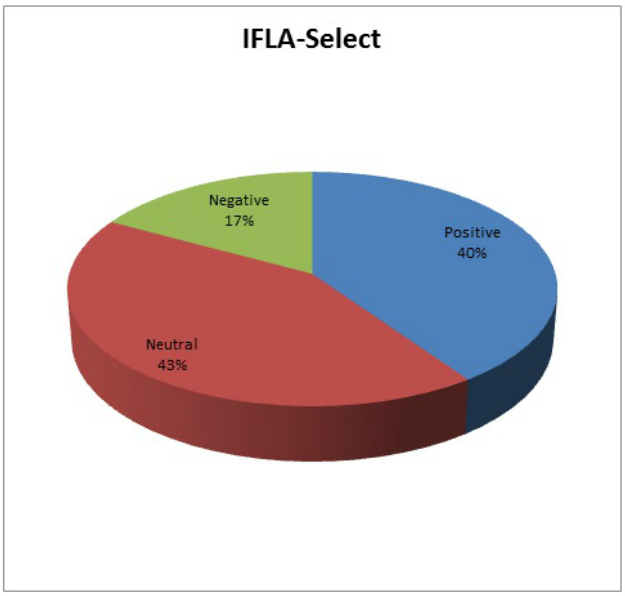


Select: Can the User Compare the Information in Multiple Records and Determine the Most Relevant Record?

At this IFLA stage, whether the relevance of a resource can be easily determined from a list with multiple records is checked. Question 13, “It is distracting to have international content listed in the results”, addresses this issue. Responses to Question 13 are summarized in Figure 7.

The analysis was based on the multilinguality parameter. Figure 7 shows that 40.0% (48/120) of evaluators are positive, meaning that they disagree that it is distracting to have international content listed in the results. However, the greatest percentage of participants (51/120, 42.5%) falls into the “neutral” category, which suggests that they have either not understood the question or do not have an opinion. Additionally, 16.7% (20/120) are negative, meaning that they agree that the multilingual content is distracting. By combining the high percentage of neutral responses with the percentage of participants who agreed with the statement (24/120, 20.0%), we come up with a high percentage of evaluators (72/120, 60.0%) that cannot possibly determine the most relevant record among multiple ones.

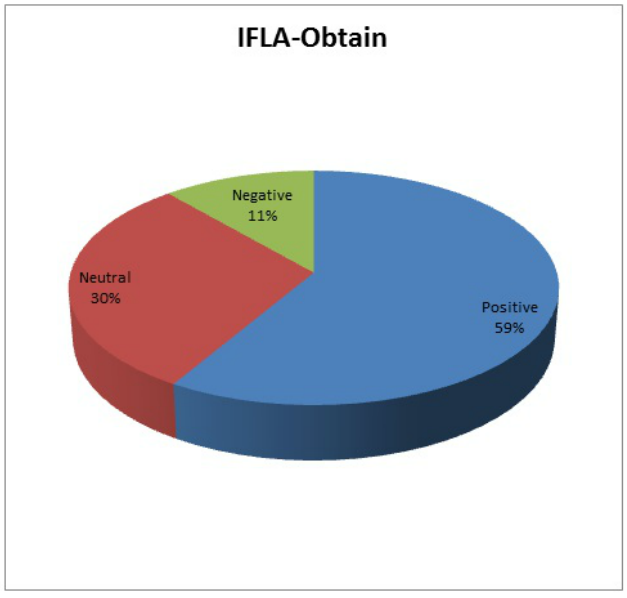
Figure 7. Users' responses mapped to the IFLA-Select requirement.



Obtain: Can the User Successfully Obtain the Original Artefacts, Based on the Information Provided in the Source Information?

In order to ascertain users' satisfaction regarding the accessibility of the original artifact, Question 14, which is "It was easy to inspect/download the retrieved learning resource", was analyzed. Responses to Question 14 are summarized in Figure 8, which shows that 59.2% (71/120) of those questioned found it easy to download the retrieved learning resource. Although more than half of the respondents replied positively, the percentage of all the rest, 40.8% (49/120), is still important and should be considered.

Figure 8. Users' responses mapped to the IFLA-Obtain requirement.



Users' Input on the Strengths and Weaknesses of the mEducator Concept

The two open-ended questions proposed at the end of the survey aimed at eliciting an overall assessment from the user of the strengths and weaknesses of the mEducator concept mostly in relation to its overall usability and value. The two open-ended questions were (1) What are the positive and negative aspects of the system for searching and finding content?, and (2) What advantages and disadvantages do you perceive of a portal with multilingual content?

As a first step, the content of the participants' answers to these two questions was analyzed for concept identification. As a second step, similar concepts with different wording were grouped into more comprehensive concepts, categorized as strengths and weaknesses. Each response was tested and re-tested for the associated concepts. Table 3 shows the details of these findings.

Table 3. Content analysis of open responses to the online survey indicating strengths and weaknesses of the mEducator system.

	Issue/Question	Count
Strengths	Awareness of international standards in digital medical education	12
	Overall, easy to use even without specific knowledge	12
	Linking with colleagues, fostering collaboration, peer reviewing	11
	Varied, interesting, up-to date content, from different providers	11
	Supports evaluation of scientific topic from different perspectives	5
	Easy access, worldwide	3
	Narrow down concepts through advanced search	3
	Can publish my own content, contribute/search in any language	3
	Specificity of information, as opposed to other search platforms	2
	Usefulness of subject profile for finding content	2
	Good concept	2
	Detailed info on terminology	1
	Is in English	1
Weaknesses	Slow search	14
	Need for translation	11
	Not user-friendly, difficult to use at the beginning	9
	Technical jargon, specific knowledge required	7
	Little variety, not enough, or irrelevant content	7
	Complex search form, issues with filtering	6
	Navigational/presentation issues	6
	Difficult to understand all content, when in foreign language	5
	Distributed semantic search difficult to understand	4
	Some items could not be accessed	3
	Some bugs	3
	No language support	3
	Most articles in English	2

The open responses provided an interesting overview of the aspects that the users spontaneously highlighted as most valuable to them. Among the strengths, the opportunities to be aware of multilingual perspectives and activities on the same topic, such as peer collaboration for research and publications purposes, were highlighted as important by users. Less frequently mentioned was the opportunity to publish own materials and the specificity of the content in the platform. Among the perceived weaknesses, recurrent themes were the slowness of the distributed search, the technical terminology that was difficult for users to understand, and the need for translation. Users' suggestions for the improvement of the mEducator system included developing a special edition of the mEducator platform to be accessed only by students, keeping the interface as simple as possible (eg, like Google), protecting against irrelevant data, providing more flexible filters for languages, and prioritizing search parameters.

Discussion

Principal Findings

This study has described how a scenario-based evaluation framework has been applied to one instantiation of the mEducator BPN (ie, mEducator3.0, MELINA+) for searching and sharing medical educational content and has presented the results of the assessment of the metadata schema used in the system and its usability and acceptance by users. The main findings of the study are that (1) the informational needs of the mEducator target groups were addressed to a satisfactory extent with regard to the process of medical content searching and sharing, (2) the metadata schema supported searching for, retrieving, and sharing of content from an end-user perspective, and (3) the overall usability of the mEducator3.0 MELINA+ system was acceptable based on the results of the SUS scores from 126 participants. Also, the study pointed out that, among the various possibilities offered by the mEducator approach to content sharing and repurposing, target users valued being aware of international standards and being able to benchmark the

resources used in other educational programs, and also being able to link with resources of their peers. In accordance with previous research [22-24], this study adopted a user-centered approach. However, one important aspect that differentiates this study from previous studies is that in the case of mEducator, completion of metadata is not performed by professional indexers [16] but is largely delegated to end-users and content providers. This is challenging because of difficulties in achieving a shared understanding of the meaning of the metadata fields. The results of this study confirmed that the meaning of metadata was not sufficiently clear to some evaluators. As reported in the facilitator's and research assistants' field notes, some specific examples where the meaning of metadata was unclear refer to descriptions of technical terms, such as "identifiers", "quality stamp", "URI", "URL", "ISSN", "ISBN", "platform users", and "external users", to name a few. Moreover, a rather high percentage of evaluators could not determine the most relevant record among multiple ones, and this could have been a result of the multilingual parameter of the system. Finally, another weakness that was documented referred to users' lack of knowledge about IPR issues, which are very important in repurposing content in medical education.

In the next section, implications for the evaluation of the processes of searching, retrieving, and sharing of data are drawn by addressing issues of content description metadata, content description procedures, and IPR for repurposed content, in the form of recommendations and lessons learned.

Recommendations and Lessons Learned From the Scenario-Based Evaluation

The scenario-based evaluation provided us with a broad and contextualized view on the user experience related to mEducator services. Repurposing and re-use of medical educational resources is becoming vital in the current economy as many of these resources are costly to produce. Systems for locating suitable medical or other content need to be improved to allow this to happen, therefore recommendations in this respect are proposed in this paper. Even though in this study, (1) medical-domain specific metadata were defined and used, (2) medical-domain–related scenarios were constructed and used, and (3) all participants were related with the medical domain, these recommendations not only apply to systems of medical education whose aim is to facilitate sharing of medical content but also apply to different systems of searching for and sharing content.

Designers of systems of searching for, retrieving, and sharing medical content should:

- emphasize ease of use, by referring to the most current best practices of usability and benchmarking their service with widely used tools (eg, Google or YouTube) and their prevailing conventions of use while simultaneously acknowledging the need for functionalities validating scientific content
- aim for smooth and effective use, by making sure the technology used does not interrupt or slow down the workflow or for example, the progressive search process
- provide support for finding and retrieving content by introducing an extensive schema in a prioritized order with the most critical metadata in front
- provide support for indexing content, that is, creating high-quality metadata, by using for example, suggestions and hints as well as dynamic, responsive annotation forms when dealing with an extensive, complex schema
- base the user interaction on the expected concepts that are familiar to the user, by simplifying the novel terminology and/or using a vocabulary with clear descriptions of technical terms (such as IPR, identifiers, quality stamp, URI, URL, ISSN, ISBN)
- support multilinguality in terms of the user experience as well as of content
- highlight the advantages, that is, the added value from using the system, for example, by offering explicit information on the type of multilingual content and system functionality available
- allow inputting metadata in various formats, especially regarding licenses or identifiers that are not compliant with users' content or are not familiar to them
- support community creation, that is, sharing contacts and content with other medical professionals, students, general practitioners, etc (this is a characteristic of mEducator and is a feature not offered by similar systems)

Based on the findings of this study, we can make several recommendations on how to improve the search, retrieval, identification, and gathering of medical resources based on the provided metadata. These recommendations not only relate to the mEducator metadata but apply to any platform that processes medical or other types of metadata to retrieve educational resources. The first recommendation is to include the option of an "advanced search". This is expected to facilitate managing retrieved results. The second recommendation is to allow different orderings of the retrieved results, for example, based on type of resource or author name. This is expected to facilitate the identification of relevant resources from among the potentially huge amount of retrieved results/records. A third recommendation is to list search terms under the results. An example would be the display of a message such as "Your search for this keyword resulted in x number of results". This feedback will enable the user to have a general idea of the quantity of retrieved results for a specific keyword or combination of keywords. Last, assessing the relevance of a resource to the user's needs before inspection can be improved in a number of different ways. One of these ways refers to the use of icons next to the list of results or also in the metadata page that could inform the users before they view the resource, that it is an "open resource" (ie, easily/freely accessible) or a resource that requires a login. Another way refers to the use of a "dashboard" in the results page, to display useful information, such as how many times the resource has been downloaded. Last, the interpretation of metadata can be improved by a thumbnail of the actual resource that should be made available while viewing the metadata.

With respect to IPR clearance, the most sensible recommendation is to devote some effort to educating the user about the options available through creative commons licensing

and easing the process of obtaining it, for example through the implementation of a step-wise process to guide the user in obtaining IPR clearance.

The users of the mEducator who want to share their content will have to specify how other users will be entitled to use it. Some recommendations derived from the evaluation efforts are the following: a better explanation of the meaning of the IPR abbreviation and options should be provided, an automated mechanism should be provided to check the validity of the IPR and prevent the user from providing any resource's link as long as the IPR had not been provided, and a policy should also be defined on how to deal with contents from external repositories for which the IPR clearance is unknown.

With regard to how the system has been adopted by the e-learning and medical community since 2011-2012 when the evaluations reported in this paper took place, it is important to note that one of the main objectives of the mEducator project was to establish standards. To this extent, the strategic decision of the consortium was to link and be involved with the activities of the Medbiquitous Consortium, an Organization producing standards for digital health education. Several fields of the mEducator schema have already been incorporated as extensions in the current process of revision of the Healthcare Learning Object Metadata (LOM) standard towards version 2.0, by the Technical Committees and LOM Working Group of Medbiquitous. Finally, it is important to mention that the mEducator3.0 MELINA+ system is under routine use in some partner organizations, but it has also been installed and trialed in other associate partners, health organizations, and IT facilities. For example, the Institute for the Study of Urological diseases is one of those using the mEducator Melina+ system.

Limitations

There were several limitations in this study, specifically, the accuracy of metadata in the stricter sense. We did not address if the metadata actually describes the educational resource that it refers to. The limited time that participants could voluntarily devote to the evaluation of mEducator as part of the different workshops, conferences, and evaluation events that were organized prohibited the use of controlled, experimental usability conditions and allowed for the administration of only two evaluation instruments and the documentation of users' difficulties through field notes. The first one of these evaluation instruments, the SUS, although generic, is a standardized instrument that was used to evaluate the mEducator3.0 MELINA+ overall usability in absolute terms, so that it can be comparable to other instantiations of the system. The second instrument was customized to address the needs of the study. Moreover, field notes taken during the evaluation sessions were used as an additional data source for triangulation purposes.

It is important to note that additional evaluation methods have been implemented in the scope of the overall study, including interviews, observational case studies, screenshot capturing, and automated tracking of activity in the different evaluation events organized by mEducator. However, these evaluation events provided space for informing initial designs and obtaining user preferences and user attitudes. The results from these different data sources were mainly used to inform the first prototype of the mEducator system in the initial stage of the project and are not reported in this study. However, final evaluations reported in this study were aiming towards more specific and important goals such as the elicitation of recommendations to the medical community.

Conclusion

The mEducator solutions (mEducator2.0 and mEducator3.0) are part of a family of tools. These tools aim to gather structured data according to the same metadata schema. The functionalities of these tools allow users maximum control over the search process, access to detailed but understandable results in order to know very quickly if the retrieved information is linked to content that is accessible or not, and the ability to browse the information in an efficient manner even if they are not entirely sure what they are looking for.

When using the mEducator services, the users' main objective is to find the resources they need in the most efficient manner. Facilitating and guiding query formulation is one of the major objectives of the mEducator tools. It is important to note that while most general purpose Web search engines deal mostly with unstructured data (and a small amount of structured data), mEducator contains almost exclusively data compliant with the mEducator metadata schema, which greatly simplifies the process of guidance in query formulation.

The different instantiations of mEducator offer somewhat different search options. Where mEducator2.0 searches only for content, the mEducator3.0 systems also allow searching for members, and some of them (eg, MetaMorphosis+ and MELINA+) allow searching for more detailed types of content such as exercises, simulations, presentations, or virtual patients. They even allow some control over the presentation or sorting of results. An important option the systems share is the option to select resource language.

The goal of this study was to implement a scenario-based evaluation methodology framework of one instantiation of the mEducator BPN (mEducator3.0/MELINA+) to assess both the metadata schema used as well as the system's overall usability and acceptance by users. The study defined the extent to which the informational needs of the mEducator target users were covered and demonstrated that the metadata schema implemented in the MELINA+ instantiation of the mEducator3.0 solution supported the searching, retrieval, and sharing of content from an end-user perspective. The study is original from a methodological point of view in that it implemented scenario-based assessment for evaluating content-sharing technologies including the interactions between end-users, content, and technology. The recommendations and lessons learned (derived with respect to content description metadata, content description procedures, and IPRs for repurposed content), the whole framework, and the results of the evaluation impact researchers, medical professionals, or other groups interested in using similar systems for educational content sharing in other domains.

Acknowledgments

This work was supported by the project mEducator (Multitype Content Sharing and Repurposing in Medical Education), funded by the eContentplus Programme, a multiannual community program to make digital content in Europe more accessible, usable, and exploitable (Grant ECP 2008 EDU 418006).

Conflicts of Interest

None declared.

Multimedia Appendix 1

SUS questionnaire.

[PDF File (Adobe PDF File), 201KB - jmir_v17i10e229_app1.pdf]

Multimedia Appendix 2

Questionnaire on the metadata and search process.

[PDF File (Adobe PDF File), 201KB - jmir_v17i10e229_app2.pdf]

References

1. Nikolaidou M, Bamidis P. ECP EDU 418006 mEducator Final Report-06-25. 2008. URL: http://www.meducator.net/sites/default/files/D9.13_mEducator_Final_Report_v.3.pdf [accessed 2014-06-25] [WebCite Cache ID 6QavfwUcH]
2. Konstantinidis S, Fernandez-Luque L, Bamidis P, Karlsen R. The role of taxonomies in social media and the semantic web for health education. A study of SNOMED CT terms in YouTube health video tags. Methods Inf Med 2013;52(2):168-179. [doi: 10.3414/ME12-02-0005] [Medline: 23450378]
3. Nicolaidou I, Bamidis P, Antoniades A, Nikolaidou M, Pattichis C, Giordano D. Multi-Type Content Sharing and Repurposing in Medical Education. 2011. p. 3740-3745 URL: http://www.editlib.org/p/38398/ [accessed 2014-06-25] [WebCite Cache ID 6c83pWbjb]
4. Dietze S, Gugliotta A, Domingue J. Supporting Interoperability and Context-Awareness in E-Learning through Situation-Driven Learning Processes. Int J Distance Educ Technol 2009;7(2):20-43. [doi: 10.4018/jdet.2009040102]
5. Bratsas C, Dimou A, Alexiadis G, Chrysou DE, Kavargyris K, Parapontis P, et al. Educational Semantic Wikis in the Linked Data Age: The Case of MSc Web Science Program at Aristotle University of Thessaloniki. 2011. URL: http://ceur-ws.org/Vol-717/paper5.pdf [accessed 2014-06-25] [WebCite Cache ID 6c7PckExO]
6. Pedrinaci C, Domingue J. Toward the Next Wave of Services: Linked Services for the Web of Data. J Univers Comput Sci 2010;16(13):1694-1719. [doi: 10.3217/jucs-016-13-1694]
7. Mitsopoulou E, Taibi D, Giordano D, Dietze S, Yu HQ, Bamidis PD, et al. Connecting medical educational resources to the Linked Data cloud: the mEducator RDF Schema, store and API. 2011. URL: http://ceur-ws.org/Vol-717/paper4.pdf [accessed 2014-06-25] [WebCite Cache ID 6QavZL5WX]
8. Giordano D, Kavasidis I, Sampinato C, Bamidis P. Developing Controlled Vocabularies for Educational Resources Sharing: A Case Study. 2011 Presented at: URL: http://ceur-ws.org/Vol-717/paper8.pdf [accessed 2014-06-25] [WebCite Cache ID 6c7Q0P7AH]
9. Bamidis P, Konstantinidis S, Bratsas C, Iyengar M. Federating learning management systems for medical education: A persuasive technologies perspective. 2011 Presented at: 24th International Symposium on Computer-Based Medical Systems; 2011; Bristol p. 1-6. [doi: 10.1109/CBMS.2011.5999117]
10. Konstantinidis S, Lakka C, Bratsas C, Pappas C, Bamidis BD. Semantic description of digital educational resources in a learning management system. 2011 Presented at: Proceedings of the 2nd Panhellenic Conference on the introduction and use of ICT in the educational process; 2011; Patras, Greece p. 413-422.
11. Dafli E, Antoniou P, Ioannidis L, Dombros N, Topps D, Bamidis PD. Virtual patients on the semantic Web: a proof-of-application study. J Med Internet Res 2015;17(1):e16 [FREE Full text] [doi: 10.2196/jmir.3933] [Medline: 25616272]
12. Kaldoudi E, Dovrolis N, Konstantinidis ST, Bamidis PD. Depicting educational content repurposing context and inheritance. IEEE Trans Inf Technol Biomed 2011 Jan;15(1):164-170. [doi: 10.1109/TITB.2010.2092442] [Medline: 21134819]
13. Kaldoudi E, Dovrolis N, Konstantinidis S, Bamidis BD. Social networking for learning object repurposing in medical education. J Inform Technol Healthc 2009;7(4):233-243.
14. Nicolaidou I, Antoniades A, Myllari J, Giordano D, Dafli E, Schizas CH, et al. Evaluation methodologies in e-systems: A proposed evaluation framework and its application in medical education. In: Bamidis P, Anastasov V, Despotova-Toleva L, editors. e-education & e-science. Bulgaria: Medical Publishing VAP; 2011:87-112.

15. Nagamori M, Sugimoto S. Using metadata schema registry as a core function to enhance usability and reusability of metadata schemas. In: DCMI International Conference on Dublin Core and Metadata Applications. 2007 Presented at: DCMI International Conference on Dublin Core and Metadata Applications; August 2007; Singapore.
16. Schwarz D, Štourač P, Komenda M, Harazim H, Kosinová M, Gregor J, et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. J Med Internet Res 2013;15(7):e135 [FREE Full text] [Medline: 23835586]
17. Rego H, Moreira T, Garcia F. AHKME Learning Object Manager Tool. In: EdMedia: World Conference on Educational Media and Technology. 2006 Presented at: EdMedia: World Conference on Educational Media and Technology; June 2006; Orlando, FL.
18. Clow D. Resource Discovery: Heavy and Light Metadata Approaches. In: EdMedia: World Conference on Educational Media and Technology. 2004 Presented at: EdMedia: World Conference on Educational Media and Technology; 2004; Lugano, Switzerland.
19. Morales-Salcedo R, Ogata H, Yano Y. Using RFID and dynamic metadata in an educational digital library. In: Proceedings of the 5th IASTED international conference on Web-based education. 2006 Presented at: 5th IASTED international conference on Web-based education; 2006; Anaheim, CA.
20. Gkatzidou S, Pearson E, Green S, Jones R. The Use of Learning Object Patterns and Metadata Vocabularies to Design Reusable and Adaptable Learning Resources. In: EDMEDIA. 2006 Presented at: EDMEDIA; 2006; Orlando, FL.
21. Zens B, Baumgartner P. Making Efficient Use of Open Educational Resources Using a Multi-Layer Metadata Approach. In: EDMEDIA. Resource Discovery; 2008 Presented at: EDMEDIA; 2008; Vienna, Austria.
22. Meire M, Ochoa X, Duval E. SAmgI: Automatic metadata generation v2.0. 2007. URL: https://lirias.kuleuven.be/handle/123456789/158919 [WebCite Cache ID 6QavuqBcE]
23. Bruce TR, Hillman DI. The Continuum of Metadata Quality: Defining, Expressing, Exploiting. 2004. URL: http://ecommons.library.cornell.edu/handle/1813/7895 [WebCite Cache ID 6Qaw67LD7]
24. Dias SB, Diniz JA, Hadjileontiadis LJ. Embracing and Embedding Techno-Pedagogical Strategies. In: Kacprzyk J, Jain LC, editors. Towards an Intelligent Learning Management System under Blended Learning. Berlin/Heidelberg: Springer Verlag; 2014:35-51.
25. Bamidis P, Kaldoudi E, Pattichis C. From taxonomies to folksonomies: a roadmap from formal to informal modeling of medical concepts and objects. In: Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine. 2009 Presented at: ITAB2009; Nov. 5-7, 2009; Larnaca, Cyprus.
26. Eysenbach G, Yihune G, Lampe K, Cross P, Brickley D. Quality management, certification and rating of health information on the Net with MedCERTAIN: using a medPICS/RDF/XML metadata structure for implementing eHealth ethics and creating trust globally. J Med Internet Res 2000;2(2 Suppl):2E1 [FREE Full text] [Medline: 11720933]
27. Heath B, McArthur D, McClelland M, Vetter R. Metadata lessons from the iLumina digital library. Commun ACM 2005 Jul 01;48(7):68-74. [doi: 10.1145/1070838.1070839]
28. Bamidis PD, Nikolaidou M, Konstantinidis S, Kontakiotis T, Triaridis S, Giordano D, et al. The burden of completing educational metadata for digital resources: a focus group study on user perceptions. In: Proceedings of the Medbiquitous Annual Conference. 2011 Presented at: Medbiquitous Annual Conference; May 7-9, 2011; Baltimore, MD.
29. Nicolaidou I, Giordano D, Spampinato C, Pattichis C, Bamidis P, Antoniades A. Evaluation of Educational Content Sharing Solutions: Challenges of the mEducator Best Practice Network. In: EDMEDIA. 2011 Presented at: EDMEDIA 2011; 2011; Lisbon, Portugal.
30. Cooper BC, Shepardson DP, Harber JM. Assessments as Teaching and Research Tools in an Environmental Problem-Solving Program for In-service Teachers. J Geosci Educ 2002;50(1):64-71.
31. Mayotte S. Online assessment of problem solving skills. Comput Human Behav 2010 Nov;26(6):1253-1258. [doi: 10.1016/j.chb.2010.03.033]
32. Looker N, Webster D, Russell D, Xu J. Scenario Based Evaluation. 2008 Presented at: Inth IEEE International Symposium on Object Oriented Real-Time Distributed Computing (ISORC); 2008; Orlando p. 2008-2011. [doi: 10.1109/ISORC.2008.56]
33. Babar M, Gorton I. Comparison of scenario-based software architecture evaluation methods. In: Software Engineering Conference, 2004. 11th Asia-Pacific; 2004:600-607.
34. Nestel D, Kneebone R, Nolan C, Akhtar K, Darzi A. Formative assessment of procedural skills: students' responses to the Objective Structured Clinical Examination and the Integrated Performance Procedural Instrument. Assessment & Evaluation in Higher Education 2011 Mar;36(2):171-183. [doi: 10.1080/02602930903221469]
35. Creutzfeldt J, Hedman L, Medin C, Heinrichs W, Felländer-Tsai L. Exploring Virtual Worlds for Scenario-Based Repeated Team Training of Cardiopulmonary Resuscitation in Medical Students. J Med Internet Res 2010;12(3):e38. [doi: 10.2196/jmir.1426]
36. Antoniou P, Athanasopoulou C, Dafli E, Bamidis P. Exploring design requirements for repurposing dental virtual patients from the web to second life: a focus group study. J Med Internet Res Jun 13 2014;16(6):e151. [doi: 10.2196/jmir.3343]
37. Luger T, Houston T, Suls J. Older Adult Experience of Online Diagnosis: Results From a Scenario-Based Think-Aloud Protocol. J Med Internet Res 2014;16(1):e16. [doi: 10.2196/jmir.2924]

38. Nicolaidou I, Antoniades A, Mylläri J, Giordano D, Dafli E, Schizas C, et al. Scenario-Based Evaluation for User Testing of Medical Educational Content-Sharing Solutions. In: INTED2012 Proceedings. 2012 Presented at: INTED2012; March 2012; Valencia, Spain p. 2619-2626.
39. Brooke J. SUS: A quick and dirty usability scale. In: Jordan P, editor. Usability evaluation in industry. London: Taylor and Francis; 1996.
40. Lewis J, Sauro J. The Factor Structure of the System Usability Scale. In: Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009. HCD. Berlin: Springer-Verlag; 2009 Presented at: 1st International Conference on Human Centered Design; 2009; Heidelberg p. 09-103. [doi: 10.1007/978-3-642-02806-9_12]
41. IFLA FRBR as User-Centered Metadata Evaluation Framework for Moving Image Collections. 2014. URL: http://www.asis.org/Conferences/AM04/posters/259.doc [accessed 2014-06-25] [WebCite Cache ID 6QawXIIFy]

Abbreviations

BPN: best practices network
CMS: content management system
CPR: cardiopulmonary resuscitation
FRBR: Functional Requirement for Bibliographic Record
HCI: human-computer interaction
LOM: learning object metadata
IFLA: International Federation of Library Associations and Institutions
IPR: intellectual property rights
ISBN: international standard book number
ISSN: international standard serial number
LCMS: learning content management system
MELINA: MEdical Education LINnked Arena
MELT: A Metadata Ecology for Learning and Teaching
SUS: System Usability Scale
URI: uniform resource identifier
URL: uniform resource locator

Edited by G Eysenbach, N Zary; submitted 26.06.14; peer-reviewed by V Waights, M Sicilia, S Kim, S Hatzipanagos, S Balula; comments to author 05.08.14; revised version received 13.10.14; accepted 19.06.15; published 09.10.15.

Please cite as:
Antoniades A, Nicolaidou I, Spachos D, Mylläri J, Giordano D, Dafli E, Mitsopoulou E, Schizas CN, Pattichis C, Nikolaidou M, Bamidis P
Medical Content Searching, Retrieving, and Sharing Over the Internet: Lessons Learned From the mEducator Through a Scenario-Based Evaluation
J Med Internet Res 2015;17(10):e229
URL: http://www.jmir.org/2015/10/e229/
doi:10.2196/jmir.3650
PMID:26453250

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org