Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 16 (2014), Issue 8   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Review

Viewpoint

Letters to the Editor

Original Paper

Interactive Sections of an Internet-Based Intervention Increase Empowerment of Chronic Back Pain Patients: Randomized Controlled Trial

Silvia Riva[1,2], PhD; Anne-Linda Camerini[1], PhD; Ahmed Allam[1], MSc; Peter J Schulz[1], PhD

[1]Institute of Communication and Health, Università della Svizzera italiana, Lugano, Switzerland

[2]Università degli Studi di Milano, Department of Health Sciences, Milan, Italy

Corresponding Author:
Silvia Riva, PhD
Institute of Communication and Health
Università della Svizzera italiana
Via Buffi 6
Lugano, 6900
Switzerland
Phone: 41 58 666 44 88
Fax: 41 58 666 46 47
Email: silvia.riva@usi.ch

Abstract

Background: Chronic back pain (CBP) represents a significant public health problem. As one of the most common causes of disability and sick leave, there is a need to develop cost-effective ways, such as Internet-based interventions, to help empower patients to manage their disease. Research has provided evidence for the effectiveness of Internet-based interventions in many fields, but it has paid little attention to the reasons why they are effective.

Objective: This study aims to assess the impact of interactive sections of an Internet-based self-management intervention on patient empowerment, their management of the disease, and, ultimately, health outcomes.

Methods: A total of 51 patients were recruited through their health care providers and randomly assigned to either an experimental group with full access to the Internet-based intervention or a control group that was denied access to the interactive sections and knew nothing thereof. The intervention took 8 weeks. A baseline, a mid-term after 4 weeks, and a final assessment after 8 weeks measured patient empowerment, physical exercise, medication misuse, and pain burden.

Results: All patients completed the study. Overall, the intervention had a moderate effect ($F_{1,52}$=2.83, P=.03, η^2=0.30, d=0.55). Compared to the control group, the availability of interactive sections significantly increased patient empowerment (midterm assessment: mean difference=+1.2, P=.03, d=0.63; final assessment: mean difference=+0.8, P=.09, d=0.44) and reduced medication misuse (midterm assessment: mean difference=−1.5, P=.04, d=0.28; final assessment: mean difference=−1.6, P=.03, d=−0.55) in the intervention group. Both the frequency of physical exercise and pain burden decreased, but to equal measures in both groups.

Conclusions: Results suggest that interactive sections as part of Internet-based interventions can positively alter patients' feelings of empowerment and help prevent medication misuse. Detrimental effects were not observed.

Trial Registration: ClinicalTrials.gov: NCT02114788; http://www.clinicaltrials.gov/ct2/show/NCT02114788 (Archived by WebCite at http://www.webcitation.org/6ROXYVoPR).

(J Med Internet Res 2014;16(8):e180) doi:10.2196/jmir.3474

KEYWORDS

Internet-based intervention; interactivity; patient empowerment; chronic back pain; health outcomes; decision; health; physical activity; pain burden; medication misuse; gamification

Introduction

Chronic Back Pain

Chronic back pain (CBP) is one of the most highly prevalent medical conditions and represents a significant public health problem. It is the second most common pain after headache and it has been identified recently as the single most important among the principal causal factors of years lived with disability worldwide [1]. As recently evaluated by Dunn et al [2], CBP is often described as a persistent condition with more than one-third of patients suffering for more than 3 years and restricting the daily activities of approximately one-third of the population annually. It is one of the most common symptomatic reasons people seek health care [2]. The costs of CBP in the European Union are considerable and have been estimated to exceed €12 billion each year [3]. As one of the most common causes of disability and sick leave, there is a need to develop new and cost-effective ways to manage the condition [2,3].

Internet-Based Interventions

One such way is through Internet-based interventions. They can play an important and compensatory role in helping CBP patients to develop appropriate self-management attitudes and strategies. In recent years, the Internet has become a prolific source for health information [4]. Today, there are hundreds, if not thousands, of health-related Internet intervention websites. In many cases, they have become a source of support for people with similar health conditions. In particular, these online programs can improve users' knowledge and perceived social support, and can therefore have a positive impact on health attitudes and the behavior of people affected by chronic conditions [5-8].

In the last 15 years, the effectiveness of Internet-based interventions has been investigated in many different chronic conditions such as headache, panic disorder, cancer, eating disorders, and, more recently, fibromyalgia and back pain [4-10]. The results of these studies are promising and indicate that Internet-based interventions are effective in improving self-management skills and self-help capabilities, and they represent a cost-effective alternative to traditional health care services [9]. The effectiveness of Internet-based interventions is now well documented by several reviews and meta-analyses [11-17].

Despite the positive outcomes of these interventions, research has also identified some limitations of assessments of Internet-based interventions [16,17]. From a methodological point of view, some findings from previous studies have been interpreted as equivocal because they did not respond to the scientific criteria of clinical trials. Most of the studies were observational, not controlled, and carried out with specific cohorts of participants [4,9]. Other studies failed to describe randomization adequately or to blind patients to the treatment group they belonged to [4,9].

From a theoretical point of view, many studies on Internet-based interventions have tried to answer the question of whether they were effective without investigating how this effect came about [18]. This means many interventions were treated as a "black box", without any noticeable focus on the different functions and components of their application. In order to understand how an intervention can be effective, the effective elements of interventions can be identified by appropriate experimental research. The design for that is straightforward: one group is given access to elements of which the effectiveness is to be studied while a control group is denied that access. The differences in measured effects can be clearly attributed to the elements under study.

Interactivity in Internet-Based Interventions

In the context of Internet-based interventions, a variable to be considered for such research is one of the major features of the Internet—its potential for interactivity [19]. There are two elements that constitute interactivity. As Sundar claimed: "One of the simplest ways to conceptualize interactivity is as a feature of the *medium*, specifically the variety of modalities that it offers for the user to experience the various parts of a website, from simple text to graphics, animation, audio and video" [20]. This variety of modalities enriches the speed, the range, and the mapping of the information, which are the three defining elements of interactivity according to Steuer [21]. *Speed* refers to "the rate at which input can be assimilated into the mediated environment", *range* refers to "the number of possibilities for action at any given time", and *mapping* refers to "the ability of a system to map its controls to changes in the mediated environment in a natural and predictable manner" [21]. The other element of interactivity is the potential for exchange. That means users cannot only choose what they get from the intervention and how they get it, they also have a chance to talk back to the medium and be talked back to in return. That is, they can ask questions and get tailored answers, they can answer questions others ask them, they can post their experience and receive reactions, and they can react to other people's experiences [22].

Interactivity can positively affect patient empowerment [19]. Patient empowerment is defined as a complex construct that includes different individual competencies and skills. According to Perkins and Zimmerman [23], empowerment goes beyond self-esteem, self-efficacy, competency, locus of control, and other traditional psychological constructs and can be considered a multilevel and multidimensional construct [24-26] closely linked to self-determination [27,28] and self-efficacy [29,30]. Moving from these considerations and favoring a psychological perspective, Thomas and Velthouse [31] proposed a cognitive model of empowerment, defined as increased intrinsic task motivation, where task motivation involves positively valued experiences that individuals derive directly from a task. In this respect, empowerment "can refer to feelings of power, control, and self-esteem that lead the patient to value autonomy—and thus interest in and desire to participate in health care decisions. This makes empowerment and its dimensions motivational constructs, and empowerment can be called volitional in this vein" [32,33]. Although evidence for the linkage between interactivity and patient empowerment is scarce [19,34], the former is said to enhance the latter because it helps individuals to be active, stimulates a positive attitude to learning, and enhances the value of autonomy [35-37]. Being a motivational construct, patient empowerment is considered a predictor of self-management behaviors, which ultimately affect the health of chronically ill patients [38].

Interactivity affects not only empowerment. Self-management behaviors as well as patients' health status are also deemed to be impacted by Internet-based interventions, especially its interactive features [39-41]. If good self-managers in reality are better able to cope with CBP, this experience should be discernable in the stories and experiences related on the interactive sections of the website [39]. A person using these sections should therefore be likely to find examples of how self-management helps other patients cope with the condition. Such positive examples should lead to the conclusion that one's own methods of self-management could be helpful in coping with the condition. This in turn should reinforce the impression that it is important what one can do to better cope and should thus, on a more general level, reinforce the impression one has of one's own empowerment in dealing with the disease. Therefore, the interactive elements in health care websites can be expected to augment health self-management.

Other outcomes that may be influenced by Internet-based interventions, particularly the interactive sections on these interventions, include physical exercise and medication adherence [40,41]. Physical exercise is generally recommended for effectively reducing or better coping with CBP [40] and is therefore the major device for self-management and, as such, a prime target of Internet-based interventions. Medication adherence (in other words, reduced medication misuse) is equally important with respect to reducing back pain without putting one's life at risk.

The ONESELF Website

This study focuses on the evaluation of a specific Internet based-intervention and its interactive features called *ONESELF* [42]. The website was first implemented in 2008 to support finding information and learning how to manage CBP and, since 2009, fibromyalgia. Research has shown that the website, which is available in Italian, was by and large successful [19,34,41,43]. It was developed by the Institute of Communication and Health of the Università della Svizzera italiana (Switzerland) in collaboration with a team of rheumatologists and physiotherapists. The health team produced the medical contents and was available to interact with subscribed patients. Communication experts reframed the contents, making them comprehensible for the general public. The website was re-launched in 2013 with a completely new interface and a widening of its scope to include rheumatic arthritis.

For this study, a modified version of the original website was created, restricting access to content on CBP only. A choice of static features including the Library, the First Aid section, and a Frequently Asked Questions (FAQ) section as well as interactive features including the Virtual Gym and the Testimonials and Commentaries sections were maintained from the ONESELF website (for a detailed description, see [19,34,41,43]). In addition, two interactive features were newly developed and implemented: a weekly Action Plan and a Quiz Game. The weekly Action Plan required patients to select at the beginning of each week from a predefined list one or more physical activities of varying intensity to be completed during the week. Reminder short message service (SMS) supported patients in complying with the plan. This feature was added based on insights into its effectiveness on chronic disease management from previous online and offline interventions [44-46]. The Quiz Game was an online examination test that allowed patients to test the information learned during navigation of the website. Patients received a multiple choice question at the end of each visited section. For every correct answer, patients earned virtual points. The sum of these points was used to classify patients in a ranking that was available to all study participants of the intervention group so that patients could see how they scored in comparison to others. This form of interactivity through feedback was proposed in the context of gamification, with the aim of using game thinking and game mechanics in non-game contexts to engage users in improving knowledge on CBP and patient empowerment [47]. Screenshots of the modified ONESELF website are available in Multimedia Appendix 1.

Study Objectives

The aim of the present randomized controlled study is to understand not only whether Internet-based interventions like ONESELF can impact patient empowerment, self-management behaviors, and, ultimately, the health status of CBP patients, but also how this can be achieved through interactive features. Thus, we propose two major hypotheses pertaining to the four desirable outcomes: patient empowerment, patients' improvement of self-management in terms of increased physical exercise and reduced medication misuse, and lower pain burden. These outcomes will improve in CBP patients over the course of the Internet-based intervention—Hypothesis 1 (H1): there will be improvement at the time of the midterm assessment over the baseline assessment and improvement at the time of the final assessment again over the baseline assessment; and Hypothesis 2 (H2): the improvement in the desirable outcomes (empowerment and physical exercise) as well as the decrease in the undesirable outcomes (medication misuse, pain burden) will be larger for CBP patients with access to the interactive sections than for patients denied this access.

Methods

Study Design

To investigate the effect of interactivity, a randomized parallel controlled study was designed (NCT02114788). Two different versions of the modified website were created, one containing only static features (ie, Library, First Aid, and FAQ) and the other containing both static and interactive features (ie, Virtual Gym, Action Plan, Testimonials and Commentaries, and Quiz Game; see Figure 1 for home page). For the intervention group, however, access to the complete version was not granted from the beginning as interactive features were added consecutively week by week as shown in Figure 2. This way, patients in the intervention group had the opportunity to become gradually familiar with the interactive features and to focus week by week on specific content and activities. Patients were blinded to the arm to which they were randomized.

Figure 1. ONESELF home page.

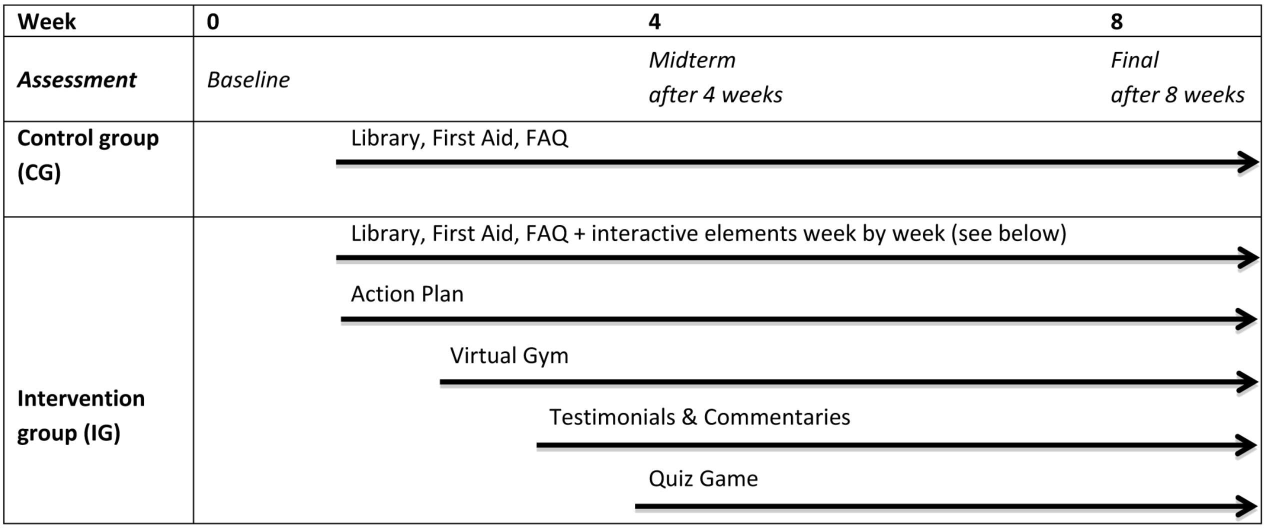
BENVENUTI IN ONESELF – IL PORTALE DELL'INFORMAZIONE PER L'AUTOGESTIONE DELLE MALATTIE REUMATICHE

Al progetto **ONESELF** collaborano, oltre agli ideatori e realizzatori del sito web, un'equipe di medici reumatologi e di fisioterapisti con i quali gli utenti possono interagire.

Il mal di schiena e' uno dei più comuni motivi di visita medica ed è il problema più frequentemente curato dopo il raffreddore.

In questa sezione riservata ai medici saranno presenti articoli selezionati dal nostro staff. E' inoltre possibile accedere ad un forum riservato.

Figure 2. Design and timeline of randomized controlled study.

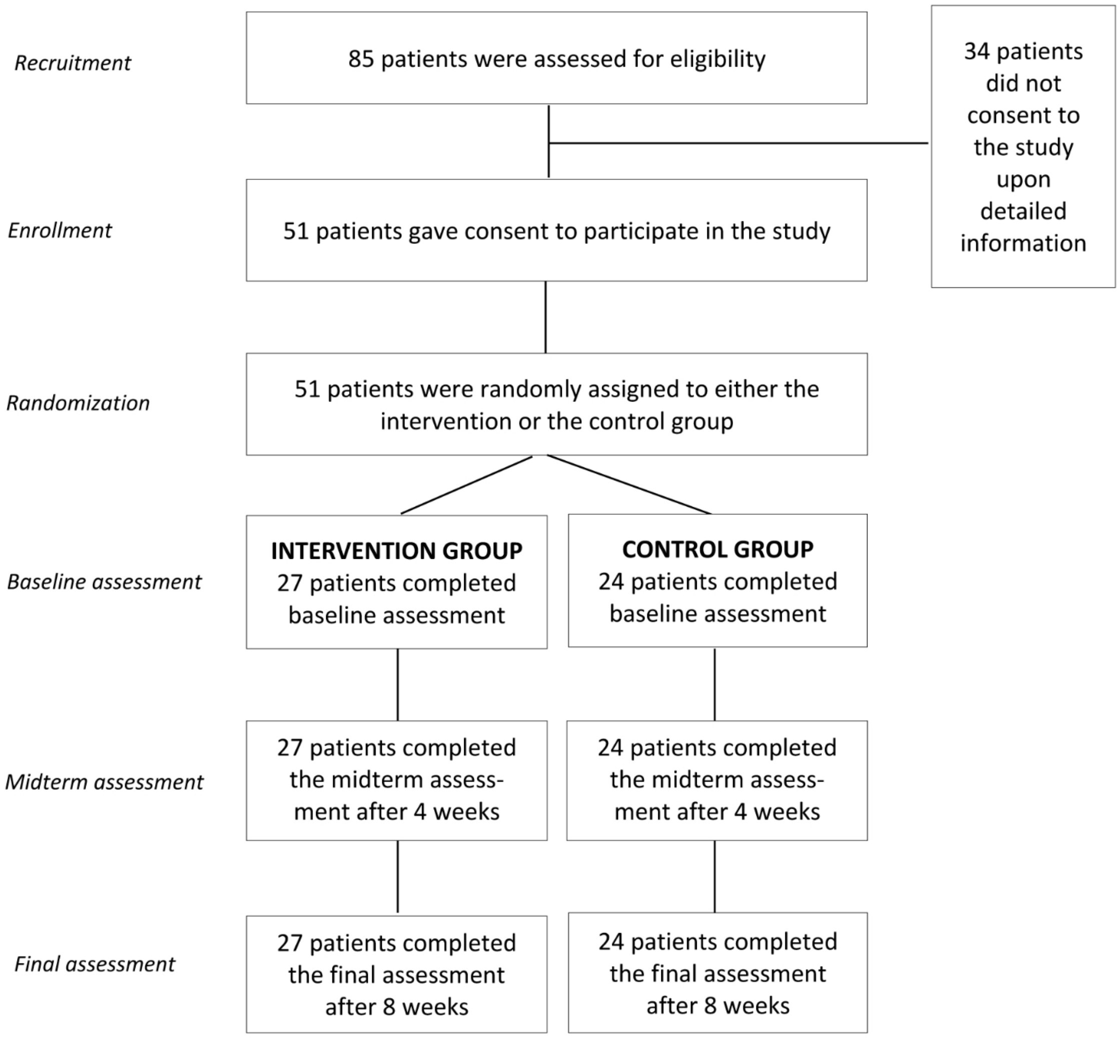


Week	**0**	**4**	**8**
Assessment	*Baseline*	*Midterm* *after 4 weeks*	*Final* *after 8 weeks*
Control group (CG)	Library, First Aid, FAQ		
Intervention group (IG)	Library, First Aid, FAQ + interactive elements week by week (see below) Action Plan Virtual Gym Testimonials & Commentaries Quiz Game		

Recruitment

Prospective participants were continually screened from February to June 2013 through their health care provider at selected clinics and rehabilitation centers in the Italian-speaking Canton Ticino (Switzerland). In each clinic and rehabilitation center, at least one health care provider was identified as a reference person who introduced the study to patients meeting predefined inclusion criteria. These were: (1) aged >18 years, (2) having suffered from back pain for at least 3 months, (3) no concurrent involvement in other studies, and (4) Italian native speakers. Patients who met these inclusion criteria and who showed interest in the study were asked to fill out a response card including their email address. A total of 85 response cards with valid email addresses were obtained. These patients subsequently received an email with a link to a detailed description of the study including an informed consent paragraph. Of the 85 interested patients, 51 eventually agreed to participate in the study. Figure 3 gives additional details of the screening, recruitment, and randomization process. The study, including the described recruitment procedure, was approved by the Ethics Committee of the Canton Ticino (Rif.CE 2337).

Figure 3. Screening, recruitment process, and random assignment.



Procedure and Randomization

The enrollment period started at the end of March 2013, and the last patient was enrolled at the end of June 2013 (3 months). Participants had access to the modified website over the course of 8 weeks. The entire study finished at the beginning of September 2013.

After confirming eligibility and obtaining informed consent from the patient (via email), the study coordinator randomly allocated participants to the two-armed parallel groups using a freely available computerized random number generator program. A permuted block randomization design method was used during the 3-month enrollment period to ensure roughly equal numbers of patients were allocated to each group. There was no face-to-face contact between the patients and research team at any point in the study, which allowed participants to live anywhere in Canton Ticino (Southern Switzerland). Of the 51 participants, 27 were allocated to the intervention group and 24 to the control group. Each participant logged in with a unique user ID so that no identifying information would be linked to their assessment, and the data were stored on secure servers. A password-protected document linking participant names to user IDs was maintained by the study coordinator, but this was not accessible to individuals involved in analyzing outcome data. Before granting access to the website, all participants were asked to complete an online questionnaire for baseline assessment. After 4 weeks, participants were asked to complete an online questionnaire for midterm assessment, and, after 8 weeks, this was repeated to get a final assessment.

Outcome Measures

Overview

All outcome measures in this study were developed and validated in English. These measures were translated into Italian and adapted to the Ticino context following standardized procedures as reported in previously published studies [19,34]. For the present study, we assessed internal consistency among patients with CBP in order to confirm the reliability of the scales in our specific context.

Empowerment

Empowerment was measured with the Psychological Empowerment Scale [31] originally developed by Thomas and

Velthouse, already cited above. The scale was originally developed for use in the workplace setting and it was adapted to be used in the health care setting [19,34]. According to the authors, empowerment is a multidimensional concept composed of four cognitive dimensions (or task assessments): (1) impact (or the degree to which behavior is seen as “making a difference”), (2) competence (or the degree to which a person can perform task activities skillfully), (3) meaningfulness (or the individual’s intrinsic caring about a given task), and (4) choice (or whether a person’s behavior is perceived as self-determined) [48]. This conceptualization aims at psychological empowerment, that is the subjective impression that one has mastery over one’s health decisions. Incorporating the multidimensionality of the concept, the scale used in this study consisted of three items adapted to the context of CBP for each of the four subdimensions: meaning (eg, “Dealing with my back pain is very important to me”), competence (eg, “I am confident about my ability to do deal with back pain”), self-determination (eg, “I have significant autonomy in determining how I deal with my back pain”), and impact (eg, “My control over the management of my back pain is large”). Participants responded on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) with higher values suggesting higher levels of psychological empowerment. At all three assessment points, the four subscales presented good internal consistency with an alpha value ranging from .71 to .94.

Medication Misuse

Medication misuse was measured with the Prescription Medication Use and Perception of Risk Instrument [49]. The scale includes six “yes/no” statements. A final sum score was obtained, which provides greater sensitivity ranging from 0 (no medication misuse) to 6 (high medication misuse) with higher scores indicating greater levels of medication misuse. Using the Kuder-Richardson-20 coefficient (KR-20) for dichotomous variables [50], psychometric testing indicated that the scale was reliable at all three assessment points (KR-20 ranging from .68 to .80).

Physical Exercise

Physical exercise in leisure time was measured with the respective subdimensions from the Short Questionnaire to Assess Health-Enhancing Physical Activity [51]. A sum score was calculated for the amount of time spent on physical exercise (hours) per week.

Pain Burden

Pain burden was measured with six items from the Chronic Pain Grading Scale [52]. Three items measured pain intensity on an 11-point scale ranging from 0 (no pain) to 10 (pain as bad as it could be). Another three items measured pain disability on an 11-point scale ranging from 0 (no interference/no change) to 10 (unable to carry on activities/extreme change). Higher values imply worse health status. At all three assessment points, the two subscales presented good internal consistency with alpha values ranging from .74 to .92 except for the pain disability scale that obtained a lower internal consistency at the final assessment (α=.62).

Data Analysis

To estimate the general effect of the intervention, between-group differences in outcome measures were analyzed with a mixed-design analysis of variance (ANOVA). A mixed-design ANOVA is used to test for differences between two or more independent groups while subjecting participants to repeated measures. Thus, in a mixed-design ANOVA model, one factor (a fixed-effects factor) is a between-subjects variable and the other is a within-subjects variable. Thus, overall, the model is a type of mixed-effect model. Between-group effect sizes were calculated according to Cohen’s *d*. Traditionally, effect sizes of 0.20 are interpreted as “small” effects, 0.50 as “moderate” effects, and 0.80 as “large” effects [53].

Testing H1, that is the improvement in patient empowerment and physical exercise as well as the decrease in medication misuse and pain burden, the development of self-report measures over the three assessment points was looked at. Changes over time were analyzed with paired samples *t* tests. Testing H2, that is the stronger improvement in patient empowerment and physical exercise as well as the stronger decrease in medication misuse and pain burden in the intervention group over the control group, made use of the randomized controlled study design and potentially yielded strong evidence for the incremental effect of interactive features over merely static informational features. Differences between the two versions of the intervention (static vs interactive) were analyzed with independent samples *t* tests and with chi-square tests for categorical variables. Differences in change over time were not determined on the aggregate but on the individual level and then averaged. This allowed the use of *t* tests for significance testing.

Eventually, a multivariate ANOVA was conducted to examine whether the conditions differed in their use of the website and the level of satisfaction with the Internet-based intervention.

Sample Size Determination

To achieve a power of 80% with 95% confidence to detect a clinically important difference of 1.0 point on the Psychological Empowerment Scale [31], assuming a standard deviation of 1.5 points similar to that found in other online intervention studies conducted in the context of ONESELF [19] and CBP [54-56], a minimum of 45 participants in total were required [57]. In the present study, 51 patients were enrolled to allow for dropouts. Calculations were done for a medium effect size (d=0.50) for group differences after the intervention.

Results

Table 1 compares participants’ sociodemographics divided by intervention and control group. There were no significant differences for any patient characteristics, although there was a trend toward higher education among participants in the intervention group.

Table 2 shows the average scores of the outcome measures studied under the different conditions and at different assessment points. More precisely, it contains mean differences in the outcome measures within the intervention group and within the control group to test the impact of the Internet-based intervention over time (H1). It furthermore contains mean differences in the outcome measures between the intervention group and the control group to test the impact of interactive elements at single assessment points (H2). No differences occurred at baseline assessment, providing support for random assignment. Although data show that at baseline assessment the control group with no access to the interactive features of the Internet-based intervention tended to feel more empowered, but to exercise less, to be less prone to medication misuse, and to experience less pain burden than the intervention group, these differences were not significant except for medication misuse.

Mean differences in the outcome measures within the control group and within the intervention group show some improvements over time, but not throughout both assessment points as H1 holds. Within the intervention group with access to the interactive website features, overall patient empowerment as a mean score of all four dimensions increased significantly at the midterm assessment (mean difference=+0.8, *P*=.01) and remained stable at the final assessment (mean difference=+0.8, *P*=.01). Among the four dimensions, the increment was higher in self-determination (mean difference=+1.7, *P*<.001), meaning (mean difference=+1.4, *P*=.03), and competence (mean difference=+1.1, *P*=.03). In contrast, within the control group without access to interactive features, no significant improvement of overall patient empowerment was evident. Furthermore, physical exercise did not improve in either of the two conditions; quite to the contrary, it declined no matter whether participants were given or denied access to the interactive features of the website and no matter whether the development up to the midterm or to the final assessment is considered. At midterm assessment, medication misuse decreased only in the intervention group with access to the interactive elements (mean difference=−0.5, *P*=.11), while it marginally significantly increased in the control group (mean difference=+1.0, *P*=.09). Only in the intervention group did the decrease continue, even if the change was not significant (mean difference=−0.6, *P*=.11).

Eventually, at final assessment after 8 weeks from the start of the intervention, pain burden significantly decreased in both conditions (control group mean difference=−1.7, *P*<.001; intervention group mean difference=−1.5, *P*<.001). The support for H1 is therefore mixed; the hypothesis draws support only from the change in pain burden and contingent upon condition, empowerment, and medication misuse, while the deterioration of physical exercise challenges the hypothesis.

With regards to H2, results of a mixed-design ANOVA show a significant difference between the two experimental conditions ($F_{1.52}$=2.83, *P*=.03, η^2=0.30, *d*=0.55). Subsequent analyses of the comparison between the two experimental conditions at midterm and at final assessment indicate that the addition of interactive features very clearly improved patients' overall empowerment. However, the majority of the differences in change from baseline assessment were significant and greater at midterm assessment (mean difference=+1.2, *P*=.03, *d*=0.63), but marginally significant and smaller at final assessment (mean difference=+0.8, *P*=.09, *d*=0.44). This is also evident considering the four dimensions of empowerment separately (especially the subdimensions meaning and self-determination). No significant difference was evident for physical exercise both at midterm and final assessment, indicating that interactivity had no incremental effect and was unable to work against the decline in exercising. For medication misuse, the differences were as expected and highly significant, meaning that interactivity clearly helped to curb this deteriorating behavior both at midterm assessment (mean difference=−1.5, *P*=.04, *d*=0.28) and final assessment (mean difference=−1.6, *P*=.03, *d*=−0.55). Eventually, interactivity had no significant effect on decrease of the burden caused by CBP. Thus, H2 receives strong support from looking at the outcomes of empowerment and medication misuse, but no support from looking at physical exercise and pain burden. There is, however, no outcome that runs against the hypothesis.

Eventually, a multivariate ANOVA was conducted to examine whether the conditions differed in their active participation in the intervention. Both the intervention and the control group were compared with regard to the use of the website, its evaluation as a means to improve CBP, and the frequency of navigation, which is the number of visited pages per week (Table 3).

Inspection of the univariate tests indicated that the difference between the two experimental conditions was significant for the frequency of navigation (*P*=.01), the evaluation of the website for improvement of CBP (*P*<.001), and the number of visited pages (*P*<.001). Participants in the intervention group, on average, used the website more often and considered it more effective for improving CBP than participants in the control group. The sections most visited by participants in the intervention group were the Library (48%, 13/27) and the Virtual Gym (33%, 9/27), while the sections more visited by the control group were the Library (80%, 19/24) and FAQ (12%, 3/24). Between-group comparison at both assessment points showed that the intervention group used the website more often and evaluated it as more beneficial. Furthermore, within- group comparison showed that in both the intervention and the control group website use significantly decreased over time showing a "wearout effect", while the evaluation of the website for CBP improvement remained stable.

Table 1. Sociodemographic characteristics of participants at baseline assessment.

Characteristic		Control group, (n=24) n (%)	Intervention group, (n=27) n (%)	Significance
Gender				0.89[a]
	Female	12 (50.0)	14 (51.9)	
	Male	12 (50.0)	13 (48.1)	
Highest educational attainment				0.13[b]
	Primary school	1 (4.1)	1 (3.7)	
	Secondary school	7 (29.1)	3 (11.1)	
	High school	13 (54.1)	14 (51.9)	
	University	3 (12.7)	9 (33.3)	
Currently in professional occupation				0.81[b]
	Yes	14 (58.3)	16 (59.3)	
	No	10 (41.7)	11 (40.7)	
Age, mean (SD)		51 (14.1)	44 (13.6)	0.58[b]
Pain duration in years, mean (SD)		9.3 (8.7)	7.9 (7.2)	0.64[b]

[a] Chi-square test

[b] Independent samples *t* test

Table 2. Means, mean differences, and significance levels for outcome measures within and between two experimental groups.

	Baseline assessment (BA)			Midterm assessment (MA)				Final assessment (FA)				Difference in change from BA to		Difference between BA and MA		Difference between BA and FA	
	CG[a]	IG[b]	Diff	CG	IG	Diff	d[c]	CG	IG	Diff	d	MA	FA	CG	IG	CG	IG
Hypothesized difference			±0			+				+		+	+	(+)	+	(+)	+
Empowerment: total score	4.5	4.0	−0.5	4.1	4.8	+0.7[d]	0.63	4.5	4.8	+0.3	0.44	+1.2[e]	+0.8[d]	−0.4	+0.8[e]	+0.05	+0.8[f]
Empowerment: meaning	4.9	4.4	−0.5	4.4	5.3	+0.9[e]	0.70	5.2	5.3	+0.1	0.09	+1.4[e]	+0.6	−0.5	+0.9[d]	+0.3	+0.9[e]
Empowerment: competence	4.5	4.0	−0.5	4.0	4.6	+0.6[d]	0.53	4.5	5.0	+0.5	−0.35	+1.1[e]	+0.9[d]	−0.5	+0.7[d]	0.0	+1.0[g]
Empowerment: self-determination	4.4	3.7	−0.7[d]	3.9	4.9	+1.0[e]	0.71	4.2	4.6	+0.4	0.27	+1.7[g]	+1.1[d]	−0.5	+1.2[g]	−0.2	+0.9[f]
Empowerment: impact	4.3	3.8	−0.5	4.2	4.6	+0.4	0.34	4.2	4.3	+0.1	0.16	+0.9	+0.6	−0.1	+0.8[d]	−0.1	+0.5
Physical exercise	1.4	2.2	+0.8	0.7	1.4	+0.7	0.36	0.3	1.3	+1.0[d]	−0.48	−0.1	+0.2	−0.7	−0.9	−1.1[e]	−0.9
Hypothesized difference			±0			−				−		−	−	(−)	−	(−)	−
Medication misuse	0.8	1.9	+1.1[e]	1.8	1.4	−0.4	0.28	2.0	1.3	−0.6	−0.55	−2.5[g]	−1.6[f]	+1.0[d]	−0.5	+1.2	−0.6
Pain burden	3.8	4.3	+0.5	3.0	3.9	+0.9[d]	0.48	2.1	2.8	+0.7	0.49	+0.4	+0.2	−0.9[d]	−0.4	−1.7[g]	−1.5[g]

[a] CG: control group

[b] IG: intervention group

[c] *d*=between-group effect sizes according to Cohen’s *d*; independent samples *t* test between CG and IG for each assessment point and for the differences in change from BA to MA and BA to FA; paired samples *t* test for differences between assessment points for control and intervention group.

[d] *P*<.10

[e] *P*<.05

[f] *P*<.01

[g] *P*<.001

Table 3. Means and mean differences for use and evaluation of the website between two experimental groups.

Use/Evaluation	Midterm assessment (MA)			Final assessment (FA)			Difference between FA and MA	
	CG[a]	IG[b]	Difference	CG	IG	Difference	CG	IG
In the last four weeks, how often did you navigate the website of the study about back pain?	1.9	2.7	+0.8[c]	1.6	2.4	+0.8[c]	−0.3	−0.3
How much has the website contributed to the improvement of your back pain in everyday life?	2.5	4.0	+1.5[e]	2.4	4.0	+1.6[e]	−0.1	0.0
Number of pages visited per week	4.0	7.0	+3.0[e]	2.8	5.0	+2.2[e]	−1.2[c]	−2.0[e]

[a] CG: control group

[b] IG: intervention group

[c] *P*<.05

[d] *P*<.01

[e] *P*<.001

Discussion

Principal Findings

Considering one of its main objectives, that is the understanding of the impact of Internet-based interventions like ONESELF on patient empowerment, the study found a moderate differential effect for the two experimental conditions. Among patients without access to the interactive sections, empowerment remained constant after 8 weeks while it significantly increased and remained consistently higher among patients who had access to the interactive sections. This suggests that the interactive sections of health care websites might indeed play a role in empowering patients with chronic conditions and gives useful insights compared to studies with contradictory results that did not pay attention to the presence or absence of interactive website features. Further evidence for the empowering effect of interactive features could be gained by looking at the actual use of these, as we would expect heavy users of interactive features to demonstrate a larger increase in empowerment than light users of these features. Future studies are needed to test this hypothesis. The differential effects of the website versions on patient empowerment refer to an overall score across all four dimensions of psychological empowerment. But they hold for each of the four dimensions too. This suggests—beyond the analyses of the psychometric qualities of this scale—that the four dimensions indeed belong together and contribute to the overall concept of empowerment. Considering the four dimensions separately, patients in the intervention group significantly improved their perceived self-determination, meaningfulness, and competence.

With regard to the differential effect of the website versions on self-management behaviors related to CBP, the results of this study show a considerable decline of physical exercise at both the midterm and the final assessment, irrespective of the experimental condition. One explanation could be that the use of websites like ONESELF, independent of the presence of interactive features, prevents people from doing what is good for them, in this case exercising to relieve pain. This, however, would run against the explicit objectives and contents of the Internet-based intervention, which put great emphasis on the necessity of exercising (the website used weekly action plans with reminder SMS messages aimed at motivating CBP patients to engage in regular physical exercise), and it would also run against the findings of other studies [58-60]. Other more probable explanations for the lack of impact on physical exercise could be a wearout and a measurement effect linked to the Internet-based intervention. The wearout effect describes the decrease in website use between the midterm and the final assessment with impact on the overall effectiveness of the intervention at final assessment. The measurement effects describes seasonal effects related to the period of enrollment since almost half of the participants (43%, 22/51) reported on their physical exercise in July and August, which are both popular holiday months in Switzerland where many people interrupt their habitual activities including physical exercise.

Results show that, overall, medication misuse did not change much as a result of the Internet-based intervention. That, however, hides very different developments in the two experimental groups: while misuse went up in the control group, it went down in the group with access to the interactive features, even though the difference between midterm and final assessment is not significant. Increased medication misuse as a consequence of a health care website is difficult to interpret but cannot be completely ruled out. No matter where the increased misuse in the control group may originate from, interactivity appears to have the potential to work against that, at least in keeping control over the use of such medications and adhering to medical regimes.

Eventually, participants experienced less pain as the exposure to the Internet-based intervention proceeded. If the intervention contributed to this decline, it was not due to its interactive features as the decrease in pain burden was observed in both groups. Strangely enough, we observed over the course of the experiment a reduction in physical exercise but a clear improvement of the pain condition. The most straightforward interpretation of this aggregate result would be that, contrary to most assumptions, the relationship between exercise and pain is different than expected. But to posit a positive relationship—more exercise, more pain—would certainly be premature, if for the fact alone that the increased misuse of medication among the control group would be difficult to explain. However, both developments could be again explained by seasonal effects. Measurement in summer might be responsible for both low levels of physical exercise due to a break of habitual behaviors for holidays, and lower levels of pain than in other times of the year with cold and rainy weather. Moreover, a lower level of back pain might be ascribed to a diminishment of work and work-related stress that can contribute to a decreased level of pain [61-64].

Back pain patients with access to the static elements of ONESELF providing information only gave up on their exercise, felt less pain, and reported more medication misuse. Participants of the intervention group with access to the interactive elements on top of the informative ones also gave up on their exercise, also felt less pain, but reported less medication misuse. These patients felt more empowered through the Internet-based intervention as compared to patients of the control group, and they reported to have better mastery over their CBP at the end of the intervention. This result was also confirmed by significant differences between the intervention group and the control group in the evaluation of the intervention as an effective means to contribute to the improvement of CBP in everyday life.

We can, therefore, conclude that the interactive features of the ONESELF website indeed contributed to improving patient empowerment while purely static elements with information only did not. Hence, this study complements the emerging literature supporting the utility of Internet-based interventions aimed at patient empowerment. The empowered patient emerges as a person who does not passively receive information, but takes increased responsibility for and a more active role in decision-making regarding his or her health [27-30]. This study highlights how empowerment is strengthened by interactivity, and this result enhances the existing literature in the field about the conjunction of these two constructs [35-37].

Limitations

This randomized controlled study is not without any limitations, which are mainly of a methodological nature. First, the study suffers from a small group size, despite the significant differences found between the two conditions. A bigger sample size might have strengthened marginally significant results and helped to detect significant differences within the intervention group for physical exercise. Second, the study lacks a pure control group. In fact, patients provided with the static version of the website were used as a control group, but no group of CBP patients was included with no access to the Internet-based intervention at all. However, the main objective of this study was to test the effectiveness of interactive sections compared to static elements only and not to test the effectiveness of the intervention as a whole. Third, a 2-month intervention might be too short a period to discover meaningful effects and conclusions on the effectiveness of Internet-based interventions on maintaining high levels of empowerment and beneficial self-management behaviors. Finally, the present study lacks specificity inasmuch as it did not take into account the quality of any of the sections that might have caused the differences between the intervention group with interactive sections and the control group with static elements only. Further insights on which specific elements cause change are essential to better inform the design of future Internet-based interventions aimed at improving chronically ill patients' empowerment, self-management behaviors, and, ultimately, their health status.

Conclusions

In conclusion, this randomized controlled study provides evidence that interactive features of Internet-based interventions aimed at chronic pain management appear mostly to affect soft outcomes related to self-perception including patient empowerment and pain representations, while the harder behavioral outcomes such as physical exercise seem to be unaffected. Nevertheless, this study adds to the growing body of literature demonstrating the effectiveness of Internet-based interventions on the management of chronic diseases like CBP. As the Internet increasingly becomes a major source of medical information and social support, this study demonstrates that it can also be an efficient and effective tool for patient empowerment that—together with health knowledge—is considered an important predictor of constructive self-management behaviors and positive health outcomes.

Acknowledgments

This project was supported by a grant from the Swiss National Science Foundation (FNSNF), Project Funding Number: FNS100013_130030/1. The authors want to thank Dr Uwe Hartung of the Institute of Communication and Health, Università della Svizzera italiana, Lugano (Switzerland) for the linguistic revision of the manuscript.

Authors' Contributions

SR participated in the design of the study, carried out the study, performed the statistical analyses, and drafted and revised the manuscript. ALC participated in the development of the design and participated in the write-up and revision of the manuscript. AA participated in the technical implementation of the intervention and the statistical analyses. PJS was the principal investigator, developed the design of the study, and participated in the write-up of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of ONESELF.

[PDF File (Adobe PDF File), 15MB - jmir_v16i8e180_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [65].

[PDF File (Adobe PDF File), 991KB - jmir_v16i8e180_app2.pdf]

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2163-2196. [doi: 10.1016/S0140-6736(12)61729-2] [Medline: 23245607]
2. Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 2013;3(12):e003838 [FREE Full text] [doi: 10.1136/bmjopen-2013-003838] [Medline: 24334157]
3. The Work Foundation. URL: http://www.theworkfoundation.com/Reports/313/The-Impact-of-Back-Pain-on-Sickness-Absence-in-Europe [accessed 2014-07-24] [WebCite Cache ID 6RJW0r8PD]
4. 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]
5. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. Patient Educ Couns 1999 Feb;36(2):193-203. [Medline: 10223023]
6. van Gemert-Pijnen JE, 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]
7. Brug J, Oenema A, Campbell M. Past, present, and future of computer-tailored nutrition education. Am J Clin Nutr 2003 Apr;77(4 Suppl):1028S-1034S [FREE Full text] [Medline: 12663313]
8. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005 Feb;29(1):54-63. [doi: 10.1207/s15324796abm2901_8] [Medline: 15677301]
9. Eysenbach G. Issues in evaluating health websites in an Internet-based randomized controlled trial. J Med Internet Res 2002 Dec;4(3):E17 [FREE Full text] [doi: 10.2196/jmir.4.3.e17] [Medline: 12554548]
10. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]
11. Murray E, Burns J, See TS, 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]
12. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
13. Ybarra ML, Eaton WW. Internet-based mental health interventions. Ment Health Serv Res 2005 Jun;7(2):75-87. [Medline: 15974154]
14. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [doi: 10.1016/j.addbeh.2005.05.002] [Medline: 15950392]
15. 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]
16. Eysenbach G. A framework for evaluating e-health: Systematic review of studies assessing the quality of health information and services for patients on the Internet. J Med Internet Res 2000 Sep 13;2(Suppl 2)(e13):1. [doi: 10.2196/jmir.2]
17. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
18. 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]
19. Camerini L, Schulz PJ. Effects of functional interactivity on patients' knowledge, empowerment, and health outcomes: an experimental model-driven evaluation of a web-based intervention. J Med Internet Res 2012;14(4):e105 [FREE Full text] [doi: 10.2196/jmir.1953] [Medline: 22810046]
20. Sundar SS. Social psychology of interactivity in human website interaction. In: Oxford Handbook of Internet Psychology. Oxford: Oxford University Press; 2007:90-102.
21. Steuer J. Defining virtual reality: dimensions determining telepresence. J Commun 1992;42(4):73-93. [doi: 10.1111/j.1460-2466.1992.tb00812.x]

22. Downes EJ, McMillan SJ. Defining interactivity: A qualitative identification of key dimensions. New Media & Society 2000 Jun 01;2(2):157-179. [doi: 10.1177/14614440022225751]
23. Perkins DD, Zimmerman MA. Empowerment theory, research, and application. Am J Community Psychol 1995 Oct;23(5):569-579. [Medline: 8851340]
24. Schulz AJ, Israel BA, Zimmerman MA, Checkoway BN. Empowerment as a multi-level construct: perceived control at the individual, organizational and community levels. Health Educ Res 1995;10(3):309-327. [doi: 10.1093/her/10.3.309]
25. Israel BA, Checkoway B, Schulz A, Zimmerman M. Health education and community empowerment: conceptualizing and measuring perceptions of individual, organizational, and community control. Health Educ Q 1994;21(2):149-170. [Medline: 8021145]
26. Zimmerman MA. Psychological empowerment: issues and illustrations. Am J Community Psychol 1995 Oct;23(5):581-599. [Medline: 8851341]
27. Deci EL. Intrinsic motivation. New York: Plenum Press; 1975.
28. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
29. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
30. Conger JA, Kanungo RN. The empowerment process: integrating theory and practice. Acad Manage Rev 1988;13:471-482.
31. Thomas KW, Velthouse BA. Cognitive elements of empowerment: an interpretive model of intrinsic task motivation. Acad Manage Rev 1990;15:666-681.
32. Schulz PJ, Nakamoto K. Emerging themes in health literacy. Stud Commun Sci 2005;5:1-10.
33. Schulz PJ, Nakamoto K. Health literacy and patient empowerment in health communication: the importance of separating conjoined twins. Patient Educ Couns 2013 Jan;90(1):4-11. [doi: 10.1016/j.pec.2012.09.006] [Medline: 23063359]
34. Camerini L, Schulz PJ, Nakamoto K. Differential effects of health knowledge and health empowerment over patients' self-management and health outcomes: a cross-sectional evaluation. Patient Educ Couns 2012 Nov;89(2):337-344. [doi: 10.1016/j.pec.2012.08.005] [Medline: 22959333]
35. Chernova S, Veloso M. Interactive policy learning through confidence-based autonomy. J Artif Intell 2009;34:1-25. [doi: 10.1613/jair.2584]
36. Wallerstein N. WHO Regional Office for Europe. Copenhagen; 2006. What is the evidence on effectiveness of empowerment to improve health? URL: http://www.gsdrc.org/go/display&type=Document&id=4052
37. Mohammad Y, Nishida T. Toward combining autonomy and interactivity for social robots. AI & Soc 2009 Feb 20;24(1):35-49. [doi: 10.1007/s00146-009-0196-3]
38. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
39. May S. Self-management of chronic low back pain and osteoarthritis. Nat Rev Rheumatol 2010 Apr;6(4):199-209. [doi: 10.1038/nrrheum.2010.26] [Medline: 20357789]
40. Hendrick P, Milosavljevic S, Hale L, Hurley DA, McDonough S, Ryan B, et al. The relationship between physical activity and low back pain outcomes: a systematic review of observational studies. Eur Spine J 2011 Mar;20(3):464-474 [FREE Full text] [doi: 10.1007/s00586-010-1616-2] [Medline: 21053026]
41. Caiata Zufferey M, Schulz PJ. Self-management of chronic low back pain: an exploration of the impact of a patient-centered website. Patient Educ Couns 2009 Oct;77(1):27-32. [doi: 10.1016/j.pec.2009.01.016] [Medline: 19321286]
42. ONESELF. URL: http://www.oneself.ch/ [accessed 2014-07-24] [WebCite Cache ID 6RJXgtGTT]
43. Camerini L, Camerini AL, Schulz PJ. Do participation and personalization matter? A model-driven evaluation of an Internet-based patient education intervention for fibromyalgia patients. Patient Educ Couns 2013 Aug;92(2):229-234. [doi: 10.1016/j.pec.2013.04.007] [Medline: 23664234]
44. Lorig K, Laurent DD, Plant K, Krishnan E, Ritter PL. The components of action planning and their associations with behavior and health outcomes. Chronic Illn 2014 Mar;10(1):50-59. [doi: 10.1177/1742395313495572] [Medline: 23838837]
45. Ahn S, Basu R, Smith ML, Jiang L, Lorig K, Whitelaw N, et al. The impact of chronic disease self-management programs: healthcare savings through a community-based intervention. BMC Public Health 2013;13:1141 [FREE Full text] [doi: 10.1186/1471-2458-13-1141] [Medline: 24314032]
46. Ory MG, Ahn S, Jiang L, Smith ML, Ritter PL, Whitelaw N, et al. Successes of a national study of the Chronic Disease Self-Management Program: meeting the triple aim of health care reform. Med Care 2013 Nov;51(11):992-998. [doi: 10.1097/MLR.0b013e3182a95dd1] [Medline: 24113813]
47. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. 2014 Presented at: 47th Hawaii International Conference on System Sciences; January 6-9, 2014; Hawaii, USA.
48. Schulz PJ, Nakamoto K. Patient behavior and the benefits of artificial intelligence: the perils of "dangerous" literacy and illusory patient empowerment. Patient Educ Couns 2013 Aug;92(2):223-228. [doi: 10.1016/j.pec.2013.05.002] [Medline: 23743214]
49. Wunsch MJ, Nakamoto K, Nuzzo PA, Behonick G, Massello W, Walsh SL. Prescription drug fatalities among women in rural Virginia: a study of medical examiner cases. J Opioid Manag 2009;5(4):228-236. [Medline: 19736903]

50. Chronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16(3):297-334.
51. Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003 Dec;56(12):1163-1169. [Medline: 14680666]
52. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992 Aug;50(2):133-149. [Medline: 1408309]
53. Cohen J. A power primer. Psychological Bulletin 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
54. Iles R, Taylor NF, Davidson M, O'Halloran P. Telephone coaching can increase activity levels for people with non-chronic low back pain: a randomised trial. J Physiother 2011;57(4):231-238 [FREE Full text] [doi: 10.1016/S1836-9553(11)70053-4] [Medline: 22093121]
55. Martorella G, Côté J, Racine M, Choinière M. Web-based nursing intervention for self-management of pain after cardiac surgery: pilot randomized controlled trial. J Med Internet Res 2012;14(6):e177 [FREE Full text] [doi: 10.2196/jmir.2070] [Medline: 23241361]
56. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]
57. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [Medline: 17695343]
58. Marangoni AH. Effects of intermittent stretching exercises at work on musculoskeletal pain associated with the use of a personal computer and the influence of media on outcomes. Work 2010;36(1):27-37. [doi: 10.3233/WOR-2010-1004] [Medline: 20555173]
59. Carpenter KM, Stoner SA, Mundt JM, Stoelb B. An online self-help CBT intervention for chronic lower back pain. Clin J Pain 2012 Jan;28(1):14-22 [FREE Full text] [doi: 10.1097/AJP.0b013e31822363db] [Medline: 21681084]
60. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;30:52-52.
61. Schulz PJ, Hartung U, Riva S. Causes, coping, and culture: a comparative survey study on representation of back pain in three Swiss language regions. PLoS One 2013;8(11):e78029 [FREE Full text] [doi: 10.1371/journal.pone.0078029] [Medline: 24223756]
62. Andersson GB. Epidemiological features of chronic low-back pain. Lancet 1999 Aug 14;354(9178):581-585. [doi: 10.1016/S0140-6736(99)01312-4] [Medline: 10470716]
63. Riva S, Monti M, Iannello P, Pravettoni G, Schulz P, Antonietti A. What does trust mean? The hidden dynamics in medical consultation. PloS ONE 2014;9(3):e90941 [FREE Full text] [doi: 10.1371/journal.pone.0090941]
64. Riva S, Monti M, Iannello P, Antonietti A. The representation of risk in routine medical experience: What actions for contemporary health policy? PloS ONE 2012;7(11):e48297 [FREE Full text] [doi: 10.1371/journal.pone.0048297]
65. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ANOVA: analysis of variance
CBP: chronic back pain
SMS: short message service

Edited by G Eysenbach; submitted 16.04.14; peer-reviewed by E Brivio, L Denti; comments to author 08.05.14; revised version received 22.05.14; accepted 31.05.14; published 13.08.14.

Please cite as:
Riva S, Camerini AL, Allam A, Schulz PJ
Interactive Sections of an Internet-Based Intervention Increase Empowerment of Chronic Back Pain Patients: Randomized Controlled Trial
J Med Internet Res 2014;16(8):e180
URL: http://www.jmir.org/2014/8/e180/
doi:10.2196/jmir.3474
PMID:25119374

Original Paper

Behavior Change Techniques Implemented in Electronic Lifestyle Activity Monitors: A Systematic Content Analysis

Elizabeth J Lyons[1,2,3], MPH, PhD; Zakkoyya H Lewis[3], BS; Brian G Mayrsohn[4], BA, MS; Jennifer L Rowland[3,5], PT, MPH, PhD

[1]The University of Texas Medical Branch, Institute for Translational Sciences, Galveston, TX, United States
[2]The University of Texas Medical Branch, Center for Interdisciplinary Research in Women's Health, Galveston, TX, United States
[3]The University of Texas Medical Branch, Department of Rehabilitation Sciences, Galveston, TX, United States
[4]University of Central Florida, College of Medicine, Orlando, FL, United States
[5]The University of Texas Medical Branch, Department of Physical Therapy, Galveston, TX, United States

Corresponding Author:
Elizabeth J Lyons, MPH, PhD
The University of Texas Medical Branch
Institute for Translational Sciences
301 University Blvd
Galveston, TX, 77555-0342
United States
Phone: 1 409 772 1917
Fax: 1 409 772 1968
Email: ellyons@utmb.edu

Abstract

Background: Electronic activity monitors (such as those manufactured by Fitbit, Jawbone, and Nike) improve on standard pedometers by providing automated feedback and interactive behavior change tools via mobile device or personal computer. These monitors are commercially popular and show promise for use in public health interventions. However, little is known about the content of their feedback applications and how individual monitors may differ from one another.

Objective: The purpose of this study was to describe the behavior change techniques implemented in commercially available electronic activity monitors.

Methods: Electronic activity monitors (N=13) were systematically identified and tested by 3 trained coders for at least 1 week each. All monitors measured lifestyle physical activity and provided feedback via an app (computer or mobile). Coding was based on a hierarchical list of 93 behavior change techniques. Further coding of potentially effective techniques and adherence to theory-based recommendations were based on findings from meta-analyses and meta-regressions in the research literature.

Results: All monitors provided tools for self-monitoring, feedback, and environmental change by definition. The next most prevalent techniques (13 out of 13 monitors) were goal-setting and emphasizing discrepancy between current and goal behavior. Review of behavioral goals, social support, social comparison, prompts/cues, rewards, and a focus on past success were found in more than half of the systems. The monitors included a range of 5-10 of 14 total techniques identified from the research literature as potentially effective. Most of the monitors included goal-setting, self-monitoring, and feedback content that closely matched recommendations from social cognitive theory.

Conclusions: Electronic activity monitors contain a wide range of behavior change techniques typically used in clinical behavioral interventions. Thus, the monitors may represent a medium by which these interventions could be translated for widespread use. This technology has broad applications for use in clinical, public health, and rehabilitation settings.

(J Med Internet Res 2014;16(8):e192) doi:10.2196/jmir.3469

KEYWORDS

electronic activity monitor; mobile; mhealth; physical activity; behavior change technique

Introduction

Background

Insufficient physical activity is a major worldwide public health problem. Even small increases in activity at a population level could have far-reaching positive impacts on chronic diseases such as diabetes, cardiovascular diseases, and several cancers [1-4]. Despite evidence supporting improved health outcomes from regular physical activity, population levels of physical activity remain low [5], and inactivity is prevalent [6].

Behavioral physical activity interventions are typically successful in increasing activity levels [7-9], but these interventions are costly and require professional expertise in delivering behavior change techniques (BCTs). Electronic activity monitors show promise as a delivery medium, as they can replicate most aspects of pedometer-based interventions while providing options for individually tailored intervention content. These monitors measure physical activity (and sometimes other health and behavior indicators such as heart rate) and interface with a computer or mobile app to provide extensive feedback tools. The feedback can be as or more rich and individualized than that provided in a clinical study, often including multiple charts, social comparisons, and indicators of progress towards individual goals. Initial intervention results using these monitors have been very promising, showing increases in physical activity and decreases in weight for two monitor brands [10-13].

The market for wearable technology activity monitors is large and growing quickly. Numerous options are currently available for use by consumers and researchers [14]. However, little is known about how these monitors differ from one another, what options they provide in their apps, and how these options may impact their effectiveness. The low cost, wide reach, and apparent effectiveness of electronic activity monitors make them appealing for recommendation by practitioners, but the growing number of options precludes practitioners' ability to provide informed recommendations to patients. Similarly, individuals interested in using a monitor to change their behavior must rely on review websites or word of mouth to compare the variety of options. Information about the functionality of the devices and the content of their companion apps could provide guidance in choosing options most similar to standard intervention practices and best suited to individual preferences and needs.

Behavior Change Techniques and Content Analyses

There is no consensus as to the best method for analyzing content of new media. A common method has been to use health behavior theory to create codes and/or percentage scores of the number of theoretical constructs represented. This process has been used with active video games [15], exercise apps [16], and weight loss apps [17].

A more involved method uses recently developed systematic taxonomies of BCTs to code for content that matches components of traditional behavioral interventions. Behavior change techniques are "observable, replicable, and irreducible component[s] of an intervention designed to alter or redirect causal processes that regulate behavior" [18]. A general hierarchical taxonomy of 93 BCTs has recently been published [18], and a similar taxonomy specific to physical activity and dietary interventions is also available [19]. These validated taxonomies likely provide a more informational and rigorous coding tool than previously used theory-based instruments.

In addition to describing the content of biomedical media, there is also a need to determine the extent to which the behavior change strategies are evidence-based. Several recent analyses have reviewed mobile apps using different sources for their evidence base. Sources have included the Expert Committee for Pediatric Obesity Prevention [20], the Health Education Curriculum Analysis Tool [21], best practices such as those used by the Diabetes Prevention Program [22,23], and clinical recommendations from the American Association of Diabetes Educators [24] and US Public Health Service [25,26].

Although no compendium of evidence-based best practices exists for exercise and weight loss behavioral interventions, several meta-analyses and meta-regressions have provided a general idea of the BCTs typically associated with successful change [27-29]. Investigating the prevalence of these techniques in particular may provide insight into future directions for research and development.

Further, determining how evidence-based BCTs are implemented may also improve standard coding methods. Preliminary evidence suggests that for several of the most common and effective techniques, fidelity to theory-based recommendations in their implementation enhances their effectiveness [30-32]. Thus, more in-depth analysis of implementation would provide valuable additional information, particularly for practitioners and those developing theory-based interventions.

Content Analysis and Electronic Activity Monitors

There has been a call for study of health apps [33-35], due to their widespread use and the absence of guidelines for determining their adherence to standard practices. In addition to standalone apps, for which there are now several published content analyses [16,17,20-22,24,25], we believe that there is a need to study electronic activity monitors and their companion apps. Development of a tool for coding the content of these apps would provide valuable information for practitioners, consumers, and researchers, and could also be used to create a decision aid for determining an appropriate match of monitoring system to individual or research/clinical intervention.

The purpose of this study was to systematically investigate currently available commercial electronic activity monitors to (1) characterize their behavior change techniques, (2) determine the extent to which they include techniques associated with successful outcomes, and (3) compare implementation of several critical techniques to theory-based and evidence-based recommendations.

Methods

Activity Monitor Inclusion Criteria and Descriptions

Monitors were included based on three sources: review listings on CNET for wearable technology, listings in the "Health and

Fitness" accessories section of the Apple Store (specifically, the Apple Web-based store that sells physical objects like iPods and MacBooks, not the online-only Apple App Store that sells apps), and a search of the Amazon site for "activity monitor". Inclusion criteria included (1) continuous monitoring of some kind of physical activity outcome (eg, continuous measurement of steps or minutes of activity rather than discrete measurement of exercise periods), (2) the most recent iteration in a similar series of products released by the same company (eg, Fitbit Force rather than Fitbit Flex), and (3) provision of feedback via a separate mobile device or personal computer interface. Additional monitors were included based on prior knowledge or suggestion of expert colleagues if they fit inclusion criteria but were not present in the three listings above (eg, Ibitz, Lumo). Descriptions of included monitors are provided in Table 1. More in-depth descriptions of each monitor with screenshots from their Web/mobile apps can be found in Multimedia Appendix 1.

Table 1. Monitor names and descriptions.

Brand	Model	Where worn	Display/ compatibility	Measures	Possible measures[a]	Food/Weight tracking
Basis		Wrist	Display, personal computer, iOS[b], Android	PA[c], Steps, Heart rate, Skin temperature, Perspiration, Sleep		
Body-Media	Fit	Upper arm	Personal computer, iOS, Android	PA, Steps, Sleep	Heart rate, Weight	Food, Weight, Balance
Fitbit	Force	Wrist	Display, personal computer, iOS, Android	PA, Steps, Sleep, Stairs, Distance, Calories	Weight	Food, Weight, Balance
Fitbug	Orb	Multiple	Personal computer, iOS, Android	Steps, Distance, Calories, Sleep	Weight	Food, Weight, Balance
Gruve		Waist	Personal computer	Calories, Activity zones		
Ibitz	Unity	Waist	iOS	Steps, Distance, Calories		Weight
Jawbone	Up24	Wrist	iOS, Android	PA, Steps, Sleep, SB[d]		Food, Weight, Balance
Lumo	Back	Waist	iOS, Android	Posture, Steps, Calories, Distance, SB, Sleep		
Misfit	Shine	Multiple	iOS, Android	Steps, Calories, Distance, PA, Sleep, "Points"		Food, Weight
Nike	Fuelband SE	Wrist	Display, personal computer, iOS, Android	PA, Steps, "Hours won", Calories, "Nikefuel"		
Polar	Loop	Wrist	Display, personal computer, iOS, Android	PA, Steps, Calories, SB, Sleep	Heart rate	
Striiv	Play	Waist	Display, personal computer, iOS, Android	PA, Steps, Stairs, Distance, Calories		Weight
Withings	Pulse	Multiple	Display, personal computer, iOS, Android	PA, Steps, Sleep, Resting heart rate	Weight, Blood pressure	Weight

[a] These objective measures are tracked simultaneously by the app. Additional measurement tools must be purchased.

[b] iOS: Apple iPhone/iPad/iPod operating system.

[c] PA: physical activity.

[d] SB: sedentary behavior.

Coding Tool and Procedure

Coding procedures for this study were based on the taxonomies of BCTs created by Michie et al [18,19]. The most recent hierarchical list was used, with published definitions guiding coding for each technique.

A tentative list of BCTs associated with successful physical activity change was created based on several recently published meta-analyses [27,28], meta-regressions [29], and systematic reviews [36-38] as well as recommendations from the US Preventive Services Task Force [39] (see Table 2).

Table 2. Behavior change techniques associated with physical activity change.

BCT #	Behavior change technique	Source
8.1	Prompt practice	[28,37,38]
2.3	Prompt self-monitoring of behavior	[29,36,37,39]
1.1	Goal-setting/intention formation	[29,36-39]
1.2	Barrier identification/problem solving	[38,39]
2.2	Provide feedback on performance	[29]
1.5	Prompt review of behavioral goals	[29]
5.1	Provide information on consequences of behavior in general	[27]
1.4	Action planning	[27]
10.3	Prompt rewards contingent on effort or progress towards behavior	[27,28]
6.2	Facilitate social comparison	[27,40]
4.1	Provide instruction	[27,36]
15.4	Self-talk	[39]
10.9	Self-rewards	[39]
3	Social support	[39]
7.1	Teach to use prompts/cues	[28]

Fidelity to implementation recommendations for three of the BCTs was measured based on Rovniak et al’s listing of recommendations for operationalizing mastery procedures from Social Cognitive Theory [30]. These are standard theory-based recommendations, but many of them also have demonstrated efficacy in randomized trials when compared to conditions that did not follow the recommendation(s) [30,32,41]. The full list of recommendations can be seen in the Results section.

Two trained coders (EL and ZL) wore each of the monitors for at least one 1-week period between November 11, 2013, and February 8, 2014. At least one coder wore each of the devices for 2 or more weeks. The coders downloaded and used personal computer apps and/or iPhone apps for each monitor. In cases where additional payment was required to access content (eg, a monthly subscription to use the BodyMedia device, a yearly subscription to access the full Fitbit website), we coded based on full access to all behavioral tools. Interrater reliability between the 2 coders was high (89%), with a kappa statistic of .55. An assistant coder also wore each monitor for 1 week and provided a full set of codes for each monitor. The 3 reviewers met to discuss any discrepancies, using the third coder’s results to help inform final decisions. The third set of codes informed final decisions in case of discrepancies. To update results, the 2 coders met once again in July 2014 to code 1 week’s worth of data on the monitors whose apps were updated since the previous data collection period. Coders also checked Web versions of apps where necessary. The same coding procedure was followed to determine whether additional techniques had been added.

Where functionality existed whereby a technique could be used but would not necessarily be used by default, we coded that technique as being present. For example, “friends” and “teams” are available for social support/social comparison in many apps, but users must add the friends themselves in order to take advantage of these tools. Further specific information on coder interpretation is available in Multimedia Appendix 2.

Results

Table 3 displays the number of monitor systems found to include each BCT. Techniques from the taxonomy that were not found in any of the systems were not included in the table. The most common techniques were those that were necessarily a part of each system: self-monitoring of behavior, feedback based on that monitoring, and the addition of a monitor to the user’s environment, behavioral goal-setting, and emphasizing a discrepancy between current behavior and goal behavior. Discrepancies were typically shown via visual progress indicators, such as progress bars, pie charts, bar charts, and line charts. Charts were often color-coded to indicate proximity to the goal, which was typically set to a default of 10,000 steps per day.

Table 3. Behavior change techniques present in monitoring systems, by number of systems (N=13).

BCT category	BCT	Monitors, n
Goals and planning	Goal setting (behavior)[a]	13
	Problem solving[a]	1
	Goal setting (outcome)	8
	Action planning[a]	5
	Review behavior goal(s)[a]	10
	Discrepancy between current behavior and goal	13
	Review outcome goal(s)	7
	Commitment	4
Feedback and monitoring	Feedback on behavior[a]	13
	Self-monitoring of behavior[a]	13
	Self-monitoring of outcome(s) of behavior	8
	Biofeedback	2
	Feedback on outcome(s) of behavior	8
Social support	Social support (unspecified)[a]	8
	Social support (practical)	2
	Social support (emotional)	4
Shaping knowledge	Instruction on how to perform the behavior[a]	2
	Information about antecedents	1
Natural consequences	Information about health consequences[a]	6
	Information about social and environmental consequences[a]	1
	Monitoring of emotional consequences	4
	Information about emotional consequences[a]	1
Comparison of behavior	Social comparison[a]	8
Associations	Prompts/cues	7
Repetition and substitution	Behavior substitution	1
	Habit formation	1
	Graded tasks	3
Comparison of outcomes	Credible source	2
Reward and threat	Non-specific reward	6
	Social reward	8
	Reward (outcome)	1
Antecedents	Adding objects to the environment	13
Scheduled consequences	Situation-specific reward	3
	Reward incompatible behavior	1
Self-belief	Focus on past successes	7

[a] This BCT was identified in the literature as associated with successful intervention.

Six techniques were present in half or more of the monitoring systems. Reviewing behavioral goals (10/13 systems) was coded when systems allowed and/or encouraged users to adjust their goals over time. Social support, social comparison, and social reward were also common (8/13 systems). Tools that allowed social support included friending systems and groups, commenting and emoticon systems for communication with others, and the ability to exercise with others virtually in real time. Social comparison was typically found in the form of lists (leaderboards), charts, and direct statements of comparison to other users. Social rewards consisted primarily of opportunities to share accomplishments and progress via social networks. Prompts or cues were found in seven systems. These were typically inactivity or idle alerts. Systems that alerted via a monitor used vibration or flashing lights to attract attention, while those that alerted via mobile device used push notifications. Seven systems also demonstrated a focus on past success, operationalized here as weekly/monthly/yearly emails discussing progress towards goals. Other techniques were found in fewer than half of the systems.

Figure 1 displays examples of screens from the Fitbit (left) and Jawbone (right) apps. See Multimedia Appendix 1 for further examples of the BCTs discussed below, each taken from one of the studied Web/mobile apps. The full listing of BCTs found in each monitoring system is presented in Multimedia Appendix 2.

Several of the more common techniques were among those found in the literature to be associated with physical activity (shown broken down by monitor system in Multimedia Appendix 3). Goal-setting, self-monitoring, and feedback were found in all of the systems. Social comparison, review of behavioral goals, social support, and social rewards were present in more than half of the monitoring systems. However, several techniques associated with successful interventions were less common. Information about consequences of the behavior and non-specific rewards were each found in six systems. Instruction on performance of the behavior, action planning, and problem solving were rare. Prompting practice, self-rewards, and self-talk were not found.

Table 4 displays theory-based recommendations for goal-setting, self-monitoring, and feedback. Overall, these recommendations were mostly followed by most of the systems. The recommendations less likely to be followed included breaking long-term goals into short-term goals (few systems included both types of goal), progression from easier to more difficult goals, tracking personally valued information, emphasizing performance successes, and comparing performance to norms of similar groups.

Table 4. Fidelity of monitoring systems to recommendations for goal-setting, self-monitoring, and feedback.

Technique	Recommendation	Monitors, n
Goal-setting	Specific	13
	Measurable	13
	Moderately challenging	13
	Long-term goals broken into short-term goals	6
	Easier goals successfully accomplished before attempting more difficult ones	3
Self-monitoring	Conducted regularly	13
	Conducted close in time to target activity	13
	Track precise information	13
	Track personally valued information	6
	Emphasize performance successes	9
	Focus on behavior modifiable by deliberate effort	13
Feedback	Specific	13
	Give a clear idea of how well participant is doing	13
	Compare performance to past accomplishments	13
	Compare performance to norms of similar groups	5
	Compare performance to precise goals	13

Figure 1. Example screen shots from Fitbit and Jawbone.

Discussion

Principal Findings

Electronic activity monitor systems include a variety of evidence-based BCTs, many of which conform to recommendations for their implementation. The most commonly found techniques were integral to the nature of the monitor: self-monitoring, feedback provision, adding objects to the environment, and goal-setting. Tools that provided or encouraged review of behavioral goals, social support, social comparison, prompts/cues, rewards, and a focus on past success were also common, found in more than half of the systems. Most of the interactive tools for goal-setting, self-monitoring, and feedback conformed to theory-based recommendations. Unfortunately, several techniques associated with successful physical activity intervention were uncommon or absent from the monitor systems, including practice, action planning, and problem solving.

Behavior Change Techniques and eHealth/mHealth

Several recently published articles have provided an overview of the current state of mobile apps for physical activity and weight loss. The results suggest that most apps do not include many BCTs that are thought to be essential to behavioral intervention. For example, one study found that 28% of pediatric obesity prevention apps included goal-setting [20], and a broader study of physical activity apps found that 44% included monitoring of some kind [21]. A study that compared weight loss apps to components of the widely used and validated Diabetes Prevention Program reported better results, with 90% or more including weight loss goals and dietary goals. However, only 20% included a physical activity goal, and fewer than 5% included problem-solving or habit formation [22]. A sample of diabetes apps were found to include a median of only two out of seven self-management behaviors recommended by the American Association of Diabetes Educators [24]. An analysis of fitness video games found greater percentages that included these techniques, such as feedback (17/18), rewards of some kind (16/18), and practice (15/18) [15]. Two content analyses of physical activity apps that used taxonomies of 23 and 26 BCTs, respectively, found that the apps included an average of 5/23 and 8/26 [42,43].

Of the activity monitor apps analyzed here, all 13 included monitoring and goal-setting. Although weight loss was not the primary purpose of most of these systems, 62% (8 /13) included weight loss goals. The prevalence of problem solving (1/13) and habit formation (1/13) was similar to that found in the other apps. As might be expected by the nature of fitness video games versus activity monitor apps, practice was much more prevalent in the video games than in the monitor systems. Rewards were fairly common in the monitor systems (6/13), but not as common as in true video games. These rewards were typical of “gamified” reward systems, including badges and achievements.

The three systems with the most techniques coded (Jawbone, Fitbit, and Nike) included 27, 20, and 19 techniques, respectively, out of the 93 possible. The absolute number of techniques found in a monitoring system may not be informative; in fact, a system with fewer but more effective techniques may ultimately produce a greater impact than a system with more numerous but less effective ones. Further, there exist several iterations of the behavior change taxonomy, ranging from 26 techniques to 93. The total number of possible techniques to be coded makes comparison across studies difficult. A recent meta-analysis of walking and cycling interventions reported that the interventions studied included a mean of approximately six BCTs, ranging from 0-12 out of a possible 26 [36]. A similar study of physical activity interventions among overweight/obese adults found 3-12 techniques out of a possible 40 [38]. The number of techniques found in the monitoring systems studied here ranged from 9-27 out of 93, with most including 12 or more. Upon recalculation using only those techniques that also exist in the 40-item taxonomy [19], we found that the activity monitor systems included 6-12 out of 40 techniques, with an average of 9 techniques per system. Recalculation using the original 26-item taxonomy found an average of 8 techniques out of 26 (range 6-12). Thus, although exact comparison is impossible due to differences in taxonomies over time, it appears that the monitors include a similar number of techniques as can be found in behavioral interventions and a potentially greater number of techniques than found in standard physical activity mobile apps.

Across multiple meta-analyses of physical activity interventions, several techniques were reported to occur in more than half of studied interventions: self-monitoring of behavior, goal-setting, providing instruction, problem solving, and prompting practice [27,28,36,44]. Self-monitoring and goal-setting were also common among the monitoring systems, but instruction, problem solving, and practice were very uncommon. Though the basic content of physical activity interventions and activity monitor systems (self-regulatory techniques such as self-monitoring, goal-setting, and feedback) are the same, monitor systems differ greatly from traditional interventions in the other implemented techniques.

Adherence to Theoretical and Empirical Best Practices

Of the 14 BCTs identified as potentially effective based on their success in previous interventions, five were widely represented across the devices: goal-setting (behavior), review of behavioral goals, feedback of behavior, self-monitoring of behavior, and rewards. Problem solving, action planning, commitment, instruction on how to perform the behavior, and behavioral practice were rare. It may be that these less common techniques are not prioritized by developers or consumers, or perhaps they are more difficult to implement. Problem solving was found in one app, but it provided only generalized tips to overcome problems (some specific to detected behavior and some discussed as common problems). Individualized problem solving would likely be a complex undertaking that would require self-report of barriers and a system for providing automated counseling. Such a system might increase app size unacceptably or be difficult to program. Action planning and commitment occurred in the context of specific challenges that users pledged or committed to undertaking. Gamification such as this appears to be a promising avenue for implementing less common BCTs.

A recent study used classification and regression trees to statistically investigate the effectiveness of combinations of BCTs [45]. The investigators found that a combination of techniques that are now called goal-setting and providing information about consequences was most successful (the analysis used a previous iteration of this taxonomy with slightly different names). They also found that interventions using feedback provision in the absence of review of behavioral goals or information about consequences were the least effective of those studied. Self-monitoring and feedback provision are the backbone of monitor systems, but many systems do not include any kind of information provision regarding specific consequences of behavior. It may be that these bare-bones apps that focus on function do not provide sufficient motivation to encourage consistent activity over time.

Several previous studies have scored apps based on their adherence to theoretical constructs. For a general study of several types of health apps, the mean score found was approximately 8/100, with the highest-scored app receiving 14/100 [17]. For a similar study specific to physical activity apps, the mean score was approximately 10/100, with the highest-scored app rated 28 [16]. It would appear that the activity monitor apps included in this analysis follow theory (here, specifically Social Cognitive Theory) more closely than apps that are not associated with activity monitors. It may be that activity monitors by definition provide behavioral tools that are suggested by Social Cognitive Theory, such as regular, instant, and precise feedback.

The above information leads to the question of what the ideal monitor and monitoring system might include. It is not surprising that two of the most well-known and popular monitors, from Fitbit and Jawbone, were highly adherent to theoretical principles (Fitbit) and evidence-based principles (Jawbone). The Jawbone Up24 was particularly impressive for including all but one of the best practice techniques investigated. However, despite their utility in previous clinical and community interventions, we do not yet know whether these techniques work well in concert and in the context of a wearable device and monitoring app.

Regardless of the number or effectiveness of BCTs included, success for an individual is likely highly influenced by individual preferences and practical issues. For example, the Misfit Shine is the only waterproof monitor of those tested and thus would likely be the most effective for someone who prefers to swim. The BodyMedia, Fitbit, Fitbug, and Jawbone systems provided energy balance information including food logs, which may make them more suitable for weight loss attempts than systems that monitored only activity and weight (although several other monitors can link to other apps that provide this service). Little is known about the reliability and validity of these devices, which could also influence user preferences. Because of the complicated series of variables that potentially influence effectiveness, a decision aid similar to those used in patient-centered outcomes research would be a logical next step for helping potential users choose a monitor in light of their preferences for techniques, game and social functions, appearance, and usability.

Clinical Applications

There exists a large and growing amount of literature demonstrating the utility of Internet- and technology-enhanced (generally called eHealth) energy balance interventions. Although several reviews have found that computer-mediated or telephone-mediated weight loss interventions were less powerful than traditional face-to-face interventions [46,47], a recent meta-regression did not find a significant effect of in-person contact on weight loss at 12 months [40]. Thus, it is currently unclear whether technology-mediated interventions can consistently replicate the effectiveness of standard clinical interventions. Activity monitor apps by definition include self-monitoring and individualized feedback, which are associated with greater effectiveness in technology-based trials [46]. These monitors may be a medium by which more effective tools can be integrated into self-directed, distance interventions.

Although little is known about the efficacy of electronic activity monitors, several clinical trials have provided preliminary data. To our knowledge, three trials have tested BodyMedia's SenseWear armband, a clinical/research grade armband that is very similar to the commercially available BodyMedia Fit armband. An early study found that adding continuous use of the armband to a 12-week standard behavioral weight loss program produced additional weight loss of approximately 2 kilograms [10]. This finding was not statistically significant in this small sample, but it may be clinically significant if distributed over a large population with respect to disease prevention and health cost reduction. A later study compared a 6-month standard behavioral weight loss (SBWL) program, SBWL plus the armband system, and the armband system alone and found a 5-kilogram difference between SBWL plus armband and SBWL [12]. A 9-month study found a 3-kilogram difference between a SBWL and SBWL plus armband group; however, this difference was not statistically significant [11].

Beyond these more typical implementations for clinical weight loss interventions, electronic activity monitors may also be a useful measure of patient-reported outcomes. Some researchers have begun using patterns of patient ambulation during and after hospitalization as a proxy measure for health, as these patterns can predict readmission [48] and other health outcomes such as quality of life and functional status [49]. Consistent, objective measures provided by these monitors could allow clinicians to identify at-risk individuals for secondary prevention and rehabilitation interventions. The CYCORE (Cyberinfrastructure for Comparative Effectiveness Research) project has demonstrated initial feasibility and acceptability of a system of home-based sensors, including activity monitors, that transmit information to oncologists for early detection of dehydration among head and neck cancer patients [50]. They could also be used to help determine appropriate lengths for hospital stays and to monitor functional independence post-release [51].

Health care professionals' preferences likely will play a role in how successful a given system is for users who are prescribed the device. Physicians, interventionists, and counselors may find that some of the companion apps are easier to integrate into their personal approach to patient care than others. Ease of surveillance may also play a role in provider choice of monitors. Some monitors more easily lend themselves to various types of surveillance, either by allowing "friends" to view user data, by partnering with other apps that allow for practitioner or friend surveillance, or by allowing users to export their data to third parties. Official methods of transmitting data to practitioners securely do not appear to currently exist in these apps. However, upcoming health information aggregator apps like those made by Apple and Google may provide a method for automatically updating physicians in the future.

Public Health and Community Applications

From a public health perspective, electronic activity monitors hold promise for large-scale, cost-effective activity and energy balance interventions. Much like previous studies of Internet-based behavioral weight loss interventions, monitor-based interventions may be less powerful than standard face-to-face programs [10,12,52]. However, they may also have a greater public health impact due to greater reach, adoption, implementation, and/or maintenance [53]. Initial investigations of the BodyMedia armband have found that it provides a more cost-effective weight loss intervention than standard behavioral weight loss interventions or combinations of the two [52].

Some of the monitors demonstrated a greater emphasis on energy balance, providing tools for monitoring intake, comparing intake to expenditure, and monitoring weight loss (eg, BodyMedia, Fitbit, Fitbug). The apps for BodyMedia, Fitbit, Fitbug, and Withings communicated with smart scales, which automatically uploaded weight measurements to the apps. These monitors and their apps provide interactive tools that mimic a large proportion of the techniques of behavioral weight loss interventions that require skilled interventionist time. These tools could reduce the time needed by interventionists for counseling by creating automated feedback.

Several of the monitoring systems included measurement and cues related to sedentary behavior. Preliminary studies that provided feedback based on baseline analyses of sedentary behavior (using research-grade monitors) produced promising results [54,55]. The existence of commercial monitors that can provide continuous real-time feedback related to sedentary behavior as well as physical activity increases the options available to interventionists. Lumo (sit time, stand ups), Polar (resting, sitting, and low intensities), and Jawbone (longest idle period) monitors measured sedentary behavior and provided mobile phone reminder alerts when sedentary periods extended past a pre-set threshold. These monitors and others that adopt this functionality could be used to implement larger-scale and lower-cost sedentary behavior interventions than those in the past [53].

Rehabilitation Applications

Electronic activity monitors have the potential to significantly improve objective measurement of physical activity for people with chronic diseases and disabilities who receive physical therapy, occupational therapy, and other types of rehabilitation services. While much has been written about the use of pedometers [56,57], accelerometers [56,58,59], and self-report questionnaires [59,60] to measure physical activity for rehabilitation patients, very little has been published expanding to other types of physical activity measurement for these populations. One exception is the development of wearable sensors, such as those described by Bonato et al [61]; however, widespread adoption and testing of wearable technology devices is not evident in the peer-reviewed literature. However, there is agreement among researchers that an effective means of quantifying physical activity is needed. Electronic activity monitors have the potential to offer a solution for gaps in current monitoring systems. For example, these monitors offer researchers and consumers the opportunity to gather physical activity data in real-world conditions such as home and community settings. They also have the capacity to provide real-world behavioral motivation using prompts and intensity measures that are variable or absent in current monitoring methods. Talkowski et al have pointed out the need for accurate physical activity intensity measures that are not currently being accurately evaluated [60]. These authors note that the number of hours of therapy is often a proxy for estimating the intensity of a rehabilitation program, whereas the length of time in therapy may not offer a uniform intensity across patients and over time. An electronic physical activity monitor would provide an objective measure of treatment intensity.

Potential for Unintended Consequences

Though numerous positive applications of these electronic activity monitors exist, there is always the possibility for unintended adverse consequences or ethical dilemmas. The potential for sharing of global positioning system (GPS) location data and personal health information produces clear privacy concerns. Surveillance of the collected data by health care providers may also lead to situations where intervention is deemed ethically necessary. Clear protocols will be necessary to guide provider behavior in such cases and to reduce risks associated with potential privacy breaches.

Because these monitors are commercially available, they can be used by individuals without consultation with medical or public health professionals. Although this widespread availability has benefits for accessibility, it might also increase the risk of negative outcomes if potentially dangerous activity programs are begun without professional oversight. The default activity goals may be inappropriate for older adults, individuals with disabilities or chronic conditions, or children [62]. Though some apps allow users to change their goals, or set goals for them based on a baseline measurement period, others provide pre-set goals that cannot be adjusted. These goals may provoke inappropriately intense activity that could lead to injury.

The validity and reliability of these monitors' step estimates is as yet unclear. Substantial literature surrounding the validity of the research grade BodyMedia armband exists (eg, [63,64]), but it is not clear whether differences between the research and commercial versions may affect energy expenditure estimates. There is preliminary evidence that one of the Fitbit monitors (worn on the waist) may produce valid estimates of steps, but distance output is inaccurate [65]. Another study found that older Fitbit monitors underestimated energy expenditure [66]. Little is known about newer, wrist-worn monitors or how monitors may differ (both from other commercial monitors and compared to gold standard measures).

Limitations

As a content analysis, this project was by definition preliminary and exploratory. Thus, our conclusions are tentative and require further study. In particular, our coding related to theory-based recommendations and our designation of specific techniques as potentially more effective than others are intended to be first steps towards formal tests as to the true impact of various recommendations or techniques. Only research with human subjects—from small qualitative investigations to large-scale randomized trials—can investigate hypotheses related to feasibility, acceptability, and effectiveness.

The systems tested here were those that measure continuous lifestyle activity that were available for purchase in late 2013. Monitors that had been discontinued (eg, Motorola MOTOACTV and Larklife) could not be tested, nor could a large number of monitors expected to be released in 2014. Follow-up tests should be conducted to include these newer monitors and compare them to earlier models. Also, we did not include monitors that focused specifically on bouts of physical activity (eg, heart rate monitors by Garmin, Polar, Mio) or mobile phone apps that measured activity using GPS or accelerometry within the phone (eg, RunKeeper, phone-based pedometer apps). Our focus on products compatible with Apple iOS, which occurred for practical reasons, may have also led to missing some monitors only available for Android devices if they were not also listed on Amazon or CNET lists. To represent the full range of available options and for use in possible future decision aids, further testing of all these monitoring systems will be necessary.

We chose to use the latest and broadest taxonomy available, which likely contributed to the greater number of techniques found in these systems. Many of the techniques in the larger taxonomy are not used in physical activity intervention (eg, many of the associations techniques are more appropriate for addiction-related interventions) and likely should not be included in activity monitor apps. It is also possible that some techniques are counterproductive or only productive in conjunction with specific other techniques. Even otherwise appropriate techniques, such as behavioral practice, may be unnecessary when the activity being promoted is an activity of daily living like walking. Which techniques are most efficacious is, of course, an empirical question not yet answered.

Finally, coding of ever-changing apps is quite difficult. Some interrater disagreement occurred because only one of the 3 testers engaged in a behavior that triggered use of a specific technique. As all monitors were tested using personal computers and iOS mobile devices, the experiences of Android users may differ from our experiences. Regular app updates also led to differential coding. Although we updated our results prior to publication, it is likely that more techniques will be included across the 13 systems and new systems will be available in the near future.

Conclusions

Electronic activity monitors include many different empirically tested behavior change techniques that are commonly implemented in clinical interventions. Many of these techniques are associated with successful physical activity and/or weight loss, and implementation of most of the techniques adhered closely to theory-based recommendations.

This content analysis provides preliminary information on the extent and type of technique implementation, thus laying a foundation for clinical, public health, and rehabilitation applications. Future studies are needed to further investigate new types of electronic activity monitors and to test their feasibility, acceptability, and ultimately their public health impact.

Acknowledgments

Dr Lyons is supported by a research career development award (K12HD052023: Building Interdisciplinary Research Careers in Women’s Health Program) from the Office of Research on Women’s Health, the Office of the Director, the National Institute of Allergy and Infectious Diseases, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development at the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dr Lyons is also supported by a Beginning Grant-in-Aid from the American Heart Association (13BGIA17110021). Dr Rowland is supported by a field-initiated grant (H133G080120) from the National Institute on Disability and Rehabilitation Research.

This study was conducted with the support of the Institute for Translational Sciences, supported in part by a Clinical and Translational Science Award (UL1TR000071) from the National Center for Advancing Translational Sciences, National Institutes of Health, and from the Sealy Center on Aging and Claude D Pepper Older Americans Independence Center (P30AG024832) at the University of Texas Medical Branch. This publication was also funded in part by the US Agency for Healthcare Research and Quality through Grant R24HS22134.

The authors would like to acknowledge the assistance of Eloisa Martinez, who served as third coder and provided additional research assistance with this project. Sedonia Holmes and Kelley Prevou also provided administrative and organizational assistance.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot examples of behavior change techniques.

[PDF File (Adobe PDF File), 1MB - jmir_v16i8e192_app1.pdf]

Multimedia Appendix 2

Behavior change techniques by monitor.

[PDF File (Adobe PDF File), 267KB - jmir_v16i8e192_app2.pdf]

Multimedia Appendix 3

Potentially effective techniques by monitor system.

[PDF File (Adobe PDF File), 74KB - jmir_v16i8e192_app3.pdf]

References

1. Sattelmair J, Pertman J, Ding EL, Kohl HW, Haskell W, Lee IM. Dose response between physical activity and risk of coronary heart disease: a meta-analysis. Circulation 2011 Aug 16;124(7):789-795 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.010710] [Medline: 21810663]
2. Woodcock J, Franco OH, Orsini N, Roberts I. Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. Int J Epidemiol 2011 Feb;40(1):121-138 [FREE Full text] [doi: 10.1093/ije/dyq104] [Medline: 20630992]
3. Friedenreich CM, Neilson HK, Lynch BM. State of the epidemiological evidence on physical activity and cancer prevention. Eur J Cancer 2010 Sep;46(14):2593-2604. [doi: 10.1016/j.ejca.2010.07.028] [Medline: 20843488]
4. Roque FR, Hernanz R, Salaices M, Briones AM. Exercise training and cardiometabolic diseases: focus on the vascular system. Curr Hypertens Rep 2013 Jun;15(3):204-214. [doi: 10.1007/s11906-013-0336-5] [Medline: 23519745]
5. Hansen BH, Kolle E, Dyrstad SM, Holme I, Anderssen SA. Accelerometer-determined physical activity in adults and older people. Med Sci Sports Exerc 2012 Feb;44(2):266-272. [doi: 10.1249/MSS.0b013e31822cb354] [Medline: 21796052]

6. Moore LV, Harris CD, Carlson SA, Kruger J, Fulton JE. Trends in no leisure-time physical activity--United States, 1988-2010. Res Q Exerc Sport 2012 Dec;83(4):587-591. [Medline: 23367822]
7. Kang M, Marshall SJ, Barreira TV, Lee JO. Effect of pedometer-based physical activity interventions: a meta-analysis. Res Q Exerc Sport 2009 Sep;80(3):648-655. [Medline: 19791652]
8. Müller-Riemenschneider F, Reinhold T, Nocon M, Willich SN. Long-term effectiveness of interventions promoting physical activity: a systematic review. Prev Med 2008 Oct;47(4):354-368. [doi: 10.1016/j.ypmed.2008.07.006] [Medline: 18675845]
9. Vaes AW, Cheung A, Atakhorrami M, Groenen MT, Amft O, Franssen FM, et al. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: A systematic review and meta-analysis. Ann Med 2013 Sep;45(5-6):397-412. [doi: 10.3109/07853890.2013.810891] [Medline: 23952917]
10. Polzien KM, Jakicic JM, Tate DF, Otto AD. The efficacy of a technology-based system in a short-term behavioral weight loss intervention. Obesity (Silver Spring) 2007 Apr;15(4):825-830. [doi: 10.1038/oby.2007.584] [Medline: 17426316]
11. Shuger SL, Barry VW, Sui X, McClain A, Hand GA, Wilcox S, et al. Electronic feedback in a diet- and physical activity-based lifestyle intervention for weight loss: a randomized controlled trial. Int J Behav Nutr Phys Act 2011;8:41 [FREE Full text] [doi: 10.1186/1479-5868-8-41] [Medline: 21592351]
12. Pellegrini CA, Verba SD, Otto AD, Helsel DL, Davis KK, Jakicic JM. The comparison of a technology-based system and an in-person behavioral weight loss intervention. Obesity (Silver Spring) 2012 Feb;20(2):356-363 [FREE Full text] [doi: 10.1038/oby.2011.13] [Medline: 21311506]
13. Barwais FA, Cuddihy TF, Tomson LM. Physical activity, sedentary behavior and total wellness changes among sedentary adults: a 4-week randomized controlled trial. Health Qual Life Outcomes 2013;11:183 [FREE Full text] [doi: 10.1186/1477-7525-11-183] [Medline: 24168638]
14. Lowe SA, Ólaighin G. Monitoring human health behaviour in one's living environment: a technological review. Med Eng Phys 2014 Feb;36(2):147-168. [doi: 10.1016/j.medengphy.2013.11.010] [Medline: 24388101]
15. Lyons EJ, Hatkevich C. Prevalence of behavior changing strategies in fitness video games: theory-based content analysis. J Med Internet Res 2013;15(5):e81 [FREE Full text] [doi: 10.2196/jmir.2403] [Medline: 23651701]
16. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
17. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
18. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
19. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
20. Schoffman DE, Turner-McGrievy G, Jones SJ, Wilcox S. Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? Transl Behav Med 2013 Sep;3(3):320-325 [FREE Full text] [doi: 10.1007/s13142-013-0206-3] [Medline: 24073184]
21. West JH, Hall PC, Hanson CL, Barnes MD, Giraud-Carrier C, Barrett J. There's an app for that: content analysis of paid health and fitness apps. J Med Internet Res 2012;14(3):e72 [FREE Full text] [doi: 10.2196/jmir.1977] [Medline: 22584372]
22. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
23. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
24. Breland JY, Yeh VM, Yu J. Adherence to evidence-based guidelines among diabetes self-management apps. Transl Behav Med 2013 Sep;3(3):277-286 [FREE Full text] [doi: 10.1007/s13142-013-0205-4] [Medline: 24073179]
25. Abroms LC, Padmanabhan N, Thaweethai L, Phillips T. iPhone apps for smoking cessation: a content analysis. Am J Prev Med 2011 Mar;40(3):279-285 [FREE Full text] [doi: 10.1016/j.amepre.2010.10.032] [Medline: 21335258]
26. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736. [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
27. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322 [FREE Full text] [doi: 10.1093/her/cyr005] [Medline: 21321008]
28. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]

29. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
30. Rovniak LS, Hovell MF, Wojcik JR, Winett RA, Martinez-Donate AP. Enhancing theoretical fidelity: an e-mail-based walking program demonstration. Am J Health Promot 2005;20(2):85-95. [Medline: 16295700]
31. Turner-McGrievy GM, Campbell MK, Tate DF, Truesdale KP, Bowling JM, Crosby L. Pounds Off Digitally study: a randomized podcasting weight-loss intervention. Am J Prev Med 2009 Oct;37(4):263-269 [FREE Full text] [doi: 10.1016/j.amepre.2009.06.010] [Medline: 19765496]
32. Pearson ES. Goal setting as a health behavior change strategy in overweight and obese adults: a systematic literature review examining intervention components. Patient Educ Couns 2012 Apr;87(1):32-42. [doi: 10.1016/j.pec.2011.07.018] [Medline: 21852063]
33. Powell AC, Landman AB, Bates DW. In search of a few good apps. JAMA 2014 May 14;311(18):1851-1852. [doi: 10.1001/jama.2014.2564] [Medline: 24664278]
34. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun;14(4):231-238. [doi: 10.1111/pedi.12034] [Medline: 23627878]
35. Goyal S, Cafazzo JA. Mobile phone health apps for diabetes management: current evidence and future developments. QJM 2013 Dec;106(12):1067-1069 [FREE Full text] [doi: 10.1093/qjmed/hct203] [Medline: 24106313]
36. Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, Powell J. Behavior change techniques used to promote walking and cycling: a systematic review. Health Psychol 2013 Aug;32(8):829-838 [FREE Full text] [doi: 10.1037/a0032078] [Medline: 23477577]
37. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
38. Bélanger-Gravel A, Godin G, Vézina-Im LA, Amireault S, Poirier P. The effect of theory-based interventions on physical activity participation among overweight/obese individuals: a systematic review. Obes Rev 2011 Jun;12(6):430-439. [doi: 10.1111/j.1467-789X.2010.00729.x] [Medline: 20331511]
39. Task Force on Community Preventive Services. Recommendations to increase physical activity in communities. Am J Prev Med 2002 May;22(4 Suppl):67-72. [Medline: 11985935]
40. Hartmann-Boyce J, Johns DJ, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. Obes Rev 2014 Jul;15(7):598-609. [doi: 10.1111/obr.12165] [Medline: 24636238]
41. Burke LE, Sereika SM, Music E, Warziski M, Styn MA, Stone A. Using instrumented paper diaries to document self-monitoring patterns in weight loss. Contemp Clin Trials 2008 Mar;29(2):182-193 [FREE Full text] [doi: 10.1016/j.cct.2007.07.004] [Medline: 17702667]
42. Middelweerd A, Mollee JS, van der Wal C, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11(1):97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
43. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
44. Ashford S, Edmunds J, French DP. What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. Br J Health Psychol 2010 May;15(Pt 2):265-288. [doi: 10.1348/135910709X461752] [Medline: 19586583]
45. Dusseldorp E, van Genugten L, van Buuren S, Verheijden MW, van Empelen P. Combinations of Techniques That Effectively Change Health Behavior: Evidence From Meta-CART Analysis. Health Psychol 2013 Nov 25:-. [doi: 10.1037/hea0000018] [Medline: 24274802]
46. 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. [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]
47. Reed VA, Schifferdecker KE, Rezaee ME, O'Connor S, Larson RJ. The effect of computers for weight loss: a systematic review and meta-analysis of randomized trials. J Gen Intern Med 2012 Jan;27(1):99-108 [FREE Full text] [doi: 10.1007/s11606-011-1803-9] [Medline: 21805218]
48. Fisher SR, Kuo YF, Sharma G, Raji MA, Kumar A, Goodwin JS, et al. Mobility after hospital discharge as a marker for 30-day readmission. J Gerontol A Biol Sci Med Sci 2013 Jul;68(7):805-810 [FREE Full text] [doi: 10.1093/gerona/gls252] [Medline: 23254776]
49. Ferriolli E, Skipworth RJ, Hendry P, Scott A, Stensteth J, Dahele M, et al. Physical activity monitoring: a responsive and meaningful patient-centered outcome for surgery, chemotherapy, or radiotherapy? J Pain Symptom Manage 2012 Jun;43(6):1025-1035. [doi: 10.1016/j.jpainsymman.2011.06.013] [Medline: 22269181]

50. Peterson SK, Shinn EH, Basen-Engquist K, Demark-Wahnefried W, Prokhorov AV, Baru C, et al. Identifying early dehydration risk with home-based sensors during radiation treatment: a feasibility study on patients with head and neck cancer. J Natl Cancer Inst Monogr 2013 Dec;2013(47):162-168. [doi: 10.1093/jncimonographs/lgt016] [Medline: 24395986]
51. Appelboom G, Yang AH, Christophe BR, Bruce EM, Slomian J, Bruyère O, et al. The promise of wearable activity sensors to define patient recovery. J Clin Neurosci 2014 Jul;21(7):1089-1093. [doi: 10.1016/j.jocn.2013.12.003] [Medline: 24534628]
52. Archer E, Groessl EJ, Sui X, McClain AC, Wilcox S, Hand GA, et al. An economic analysis of traditional and technology-based approaches to weight loss. Am J Prev Med 2012 Aug;43(2):176-182. [doi: 10.1016/j.amepre.2012.04.018] [Medline: 22813682]
53. Mayrsohn B, Khalil G. Games for Health. In: Krohn R, Metcalf D, editors. mHealth Innovation: Best Practices from the Mobile Frontier. Chicago, IL: HIMSS; 2014:85-94.
54. Fitzsimons CF, Kirk A, Baker G, Michie F, Kane C, Mutrie N. Using an individualised consultation and activPAL™ feedback to reduce sedentary time in older Scottish adults: results of a feasibility and pilot study. Prev Med 2013 Nov;57(5):718-720. [doi: 10.1016/j.ypmed.2013.07.017] [Medline: 23891853]
55. Gardiner PA, Eakin EG, Healy GN, Owen N. Feasibility of reducing older adults' sedentary time. Am J Prev Med 2011 Aug;41(2):174-177. [doi: 10.1016/j.amepre.2011.03.020] [Medline: 21767725]
56. Berlin JE, Storti KL, Brach JS. Using activity monitors to measure physical activity in free-living conditions. Phys Ther 2006 Aug;86(8):1137-1145 [FREE Full text] [Medline: 16879047]
57. Coleman KL, Smith DG, Boone DA, Joseph AW, del Aguila MA. Step activity monitor: long-term, continuous recording of ambulatory function. J Rehabil Res Dev 1999 Jan;36(1):8-18. [Medline: 10659890]
58. Murphy SL. Review of physical activity measurement using accelerometers in older adults: considerations for research design and conduct. Prev Med 2009 Feb;48(2):108-114. [doi: 10.1016/j.ypmed.2008.12.001] [Medline: 19111780]
59. Henderson CJ, Lovell DJ, Specker BL, Campaigne BN. Physical activity in children with juvenile rheumatoid arthritis: quantification and evaluation. Arthritis Care Res 1995 Jun;8(2):114-119. [Medline: 7794985]
60. Talkowski JB, Lenze EJ, Munin MC, Harrison C, Brach JS. Patient participation and physical activity during rehabilitation and future functional outcomes in patients after hip fracture. Arch Phys Med Rehabil 2009 Apr;90(4):618-622. [doi: 10.1016/j.apmr.2008.10.024] [Medline: 19345777]
61. Bonato P. Advances in wearable technology and applications in physical medicine and rehabilitation. J Neuroeng Rehabil 2005 Feb 25;2(1):2 [FREE Full text] [doi: 10.1186/1743-0003-2-2] [Medline: 15733322]
62. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De Bourdeaudhuij I, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011;8:80 [FREE Full text] [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
63. Berntsen S, Hageberg R, Aandstad A, Mowinckel P, Anderssen SA, Carlsen KH, et al. Validity of physical activity monitors in adults participating in free-living activities. Br J Sports Med 2010 Jul;44(9):657-664. [doi: 10.1136/bjsm.2008.048868] [Medline: 18628358]
64. Cole PJ, LeMura LM, Klinger TA, Strohecker K, McConnell TR. Measuring energy expenditure in cardiac patients using the Body Media Armband versus indirect calorimetry. A validation study. J Sports Med Phys Fitness 2004 Sep;44(3):262-271. [Medline: 15756165]
65. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2013 Oct 31:-. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
66. Dannecker KL, Sazonova NA, Melanson EL, Sazonov ES, Browning RC. A comparison of energy expenditure estimation of several physical activity monitors. Med Sci Sports Exerc 2013 Nov;45(11):2105-2112. [doi: 10.1249/MSS.0b013e318299d2eb] [Medline: 23669877]

Abbreviations

BCTs: behavior change techniques
GPS: global positioning systems
iOS: Apple iPhone/iPad/iPod operating system
SBWL: standard behavioral weight loss

Edited by G Eysenbach; submitted 14.04.14; peer-reviewed by C Pellegrini, W Evans, A Tallner; comments to author 16.07.14; revised version received 30.07.14; accepted 30.07.14; published 15.08.14.

Please cite as:
Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL
Behavior Change Techniques Implemented in Electronic Lifestyle Activity Monitors: A Systematic Content Analysis
J Med Internet Res 2014;16(8):e192
URL: http://www.jmir.org/2014/8/e192/
doi:10.2196/jmir.3469
PMID:25131661

Original Paper

How an Online Intervention to Prevent Excessive Gestational Weight Gain Is Used and by Whom: A Randomized Controlled Process Evaluation

Margaret Mochon Demment[1], PhD; Meredith Leigh Graham[1], MS; Christine Marie Olson[1], PhD

Community Nutrition, Division of Nutritional Sciences, Cornell University, Ithaca, NY, United States

Corresponding Author:
Margaret Mochon Demment, PhD
Community Nutrition
Division of Nutritional Sciences
Cornell University
370 MVR Hall
Ithaca, NY, 14853
United States
Phone: 1 607 255 2634
Fax: 1 607 255 1033
Email: mmd238@cornell.edu

Abstract

Background: Online interventions have emerged as a popular strategy to promote healthy behaviors. Currently, there is little agreement about how to capture online intervention engagement. It is also uncertain who engages with weight-related online interventions and how engagement differs by demographic and weight characteristics.

Objective: The objectives of this study were to (1) characterize how pregnant women engaged with features of an online intervention to prevent excessive gestational weight gain, (2) identify demographic and weight status subgroups of women within the sample, and (3) examine differences in use of intervention features across the demographic and weight status subgroups.

Methods: A sample of racially and socioeconomically diverse pregnant women from a northeastern US city was assigned to the intervention group in a randomized controlled trial to prevent excessive gestational weight gain (n=1014). The intervention website included these features: weight-gain tracker, health-related articles, blogs, physical activity and diet goal-setting tools, and local resources. Engagement variables were created to capture the amount, consistency, and patterns of feature use across pregnancy using latent class analysis. Demographic/weight status subgroups were also created using latent class analysis. Differences in engagement across the demographic/weight status subgroups were examined using chi-square analysis.

Results: Six engagement patterns emerged: “super-users” (13.02%, 132/1014), “medium-users” (14.00%, 142/1014), “consistent weight-tracker users” (14.99%, 152/1014); “almost consistent weight-tracker users” (21.99%, 223/1014), “inconsistent weight-tracker users” (15.98%, 162/1014), and “non-users” (20.02%, 203/1014). Four demographic/weight status subgroups emerged: three minority and one white. There were different engagement patterns by demographic/weight status subgroups. Super-users were more likely to be in the white subgroup, while non-users were more likely to be in the minority subgroups. However, around a third of women in minority subgroups were consistently or almost consistently engaging with the weight-tracker (black, young women, 32.2%, 67/208; black, heavier women, 37.9%, 50/132; Hispanic women, 27.4%, 32/117).

Conclusions: While white and higher income women had higher engagement in general, depending on the measure, there was still considerable engagement by the minority and low-income women.

Trial Registration: Clinicaltrials.gov: NCT01331564; http://clinicaltrials.gov/ct2/show/NCT01331564 (Archived by WebCite at http://www.webcitation.org/6Rw4yKxI5).

(J Med Internet Res 2014;16(8):e194) doi:10.2196/jmir.3483

KEYWORDS

online intervention; obesity prevention; latent class analysis; socioeconomic differences; demographic subgroups; online engagement; process evaluation

Introduction

Online interventions have emerged as a popular strategy in obesity prevention. In order to understand the effectiveness of these interventions, it is critical to identify how participants use online interventions and if there are differences by demographic characteristics and weight status. Process evaluation is an important step for understanding how interventions achieve their intended outcomes. A key process measure is dose of intervention received. A higher dose of intervention received or higher engagement has generally been associated with greater success in achieving weight-related intervention outcomes [1-3]. This has been particularly true in online behavior change interventions [4-6]. While online interventions provide a unique opportunity to measure engagement objectively, there is no current consensus on the definitions and measures for engagement in online interventions [7]. Previous studies have used the following measures of use: number of website visits or log-ins; time spent on a site; and number of features used [7-9]. No studies to date have examined feature use based on expected use, consistency of use across the intervention time period, or how usage clusters across different intervention features.

In this study, online engagement with a website to prevent excessive gestational weight gain (GWG) is examined. Avoiding excessive GWG, defined as gaining more than the Institute of Medicine's GWG recommendations, has become a priority in obesity prevention [10]. Excessive GWG is a risk factor for postpartum weight retention that contributes to long-term weight gain [11]. Low-income and minority women are more likely to be overweight and obese and are more likely to gain more weight than recommended during pregnancy [12]. Given this increased risk for excessive GWG, it would be valuable to know how low-income and minority women engage with an online intervention. Differences in engagement for GWG interventions by demographic characteristics have not been examined. However, rates of participation in diabetes self-management trials were highly variable in diverse samples inclusive of lower income individuals and minorities [13]. In addition, in a 2-year weight loss trial, obese non-white participants were significantly less likely to self-weigh weekly compared to white participants [13,14].

In most studies, demographic and weight status variables are examined independently and statistical methods isolate the independent effect of each factor. This has several methodological challenges as articulated by Lanza et al [15] and summarized here: (1) it can lead to Type 1 error (eg, the need for multiple comparisons to be run on each characteristic increasing the likelihood of finding a significant result), (2) the statistical power to detect an effect varies by characteristics (depending on the number of individuals within the categories for the characteristic), and (3) higher order interactions are often impossible to evaluate due to sample size constraints (eg, comparing older white females to younger white females) [16]. For this study, we use an alternative method that has emerged in prevention medicine [16-18]. Rather than focusing on isolated risk factors or characteristics, individuals with all their demographic characteristics and weight status are first categorized into multidimensional subgroups through latent class analysis and then the subgroup variable is used in subsequent analyses, rather than independent variables for each demographic characteristic and weight status.

This study addresses two gaps in the current understanding of weight-related online interventions. First, this study provides new measures of engagement that consider expected use, consistency of use across time, and patterns of use for different features. Second, this study examines how multiple measures of engagement differ across demographic/weight status subgroups. The aims of this study were to (1) characterize how pregnant women engaged with online intervention features, (2) identify demographic and weight status subgroups of women within the sample, and (3) examine differences in use of intervention features across the demographic and weight status subgroups.

Methods

Intervention

Fishbein and Yzer's Integrative Model of Behavioral Prediction [19] was the guiding theoretical framework for the online intervention to prevent excessive gestational weight gain. Fishbein and Yzer's framework was combined with Fogg's Behavior Model for Persuasive Design [20] to link behaviors and their predictors to intervention features. The online intervention included blogs, local resources, articles, frequently asked questions (FAQs), and events. In addition, intervention participants also had access to the weight gain tracker and diet and physical activity goal-setting tools. Intervention participants were emailed weekly from randomization to delivery with new content and reminders to use the weight gain tracker, diet, and physical activity goal-setting tools. Intervention features are described in more detail in Graham et al 2014 [21] and images are available in Multimedia Appendix 1.

Participants

Data from the intervention group of a randomized controlled trial of excessive GWG and postpartum weight retention prevention with women who were 18-35 years of age, normal to obese class I body mass index (BMI), socially and racially diverse, and relatively healthy (N=1689), conducted in the northeastern United States, were used in this study. That randomized trial, conducted from 2011-2014, is described in detail elsewhere [22,23]. To meet eligibility criteria, participants had to (1) consent at or before 20 weeks gestation, (2) be available for a 24-month intervention, (3) plan to carry their pregnancy to term and keep their baby, (4) read and understand English, and (5) have an email address. Exclusion criteria included: BMI<18.5 kg/m^2(underweight) or >35.0 kg/m^2(class 2 obese), multiple gestation (eg, twins), having had eating disorders or gastric bypass surgery in the past, having had three or more consecutive miscarriages, and the presence of pre-pregnancy medical conditions that could influence weight loss or gain. All study participants were sent an email describing the tools on the website, and email, a postcard, and telephone reminders were used as prompts to encourage participants to visit the website the first time. A US$5 incentive was also given for the first website visit. There were 1126 eligible women who were assigned to the intervention group for pregnancy. The sample for this analysis included women who participated in the study through pregnancy (ie, did not withdraw, miscarry, or have a pre-term birth at less than 28 weeks gestation) (n=1014). For this analytic sample, participants were exposed to the intervention for a minimum of 2 months and a maximum of 9 months depending on both week gestation at enrollment and week gestation at delivery. The women excluded from analysis did not significantly differ from those included (Table 1).

Table 1. Sample characteristics.

Characteristic		Total sample, n=1689 n (%)	Intervention sample, n=1126 n (%)	Analysis sample, n=1014 n (%)	*P* value[a]
Race					.28
	White	1054 (62.40)	693 (61.55)	630 (62.13)	
	Black	395 (23.39)	273 (24.25)	239 (23.57)	
	Other	240 (14.21)	160 (14.21)	145 (14.30)	
Hispanic					.44
	Yes	212 (12.55)	145 (12.88)	128 (12.62)	
	No	1477 (87.45)	981 (87.12)	886 (87.38)	
Low-income					.57
	Yes	734 (43.46)	494 (43.87)	442 (42.59)	
	No	955 (56.54)	632 (56.13)	572 (56.41)	
Body mass index category					.74
	Normal	872 (51.63)	575 (51.07)	520 (51.28)	
	Overweight	508 (30.08)	346 (30.73)	308 (30.37)	
	Obese	309 (18.29)	205 (18.21)	186 (18.34)	
Age category, years					.90
	18 - <25	506 (29.96)	341 (30.28)	305 (30.08)	
	25 - <30	546 (32.33)	363 (32.24)	328 (32.35)	
	≥30 years	637 (37.71)	422 (37.48)	381 (37.57)	

[a] Chi-square analysis *P* value comparing analysis sample and those not included (n=112) from the intervention sample.

Data Collection

Five sources of data were used in this analysis: screening for eligibility, postpartum height and weight visit, medical chart audit, website activity, and survey. At baseline screening, which took place at less than 20 weeks gestation, the following self-reported variables were collected: race, ethnicity, date of birth, height, current weight, early pregnancy (<13 weeks) or pre-pregnancy weight, and a measure of low-income using a participant's insurance type to determine if a participant qualified for Women, Infants and Children Program (WIC)/ Medicaid/ Prenatal Care Assistance Program (PCAP). The self-reported ethnicity question asked, "Are you of Hispanic or Latino origin?" with the response categories of "yes" and "no". The self-reported race item asked, "Which race best describes you?" with the response categories of "American Indian and Alaska Native", "Asian", "Black or African American", "White", "Native Hawaiian and other Pacific Islander", and "Other race, please specify".

To categorize women's weight status, measured weight and height were used to calculate BMI for the vast majority of the sample. Height was collected from three data sources: (1) measured height from postpartum weight collection visits (1307/1689, 77.38% of sample), (2) measured height from the medical chart (358/1689, 21.20% of sample), or (3) self-reported height at screening (24/1689, 1.42% of sample). Pre-pregnancy or early pregnancy weight was collected from three data sources: (1) measured early pregnancy weight from the medical chart (1599/1689, 94.67%), (2) self-reported or measured pre-pregnancy weight from the medical chart (67/1689, 3.97%), or (3) self-reported pre-pregnancy or early pregnancy weight at screening (23/1689, 1.36% of sample).

Medical chart audit data were used to verify and correct the date of birth of the participant (33/1689 individuals with changed date of births, 1.95% of total study sample). Date of birth and date of consent were used to calculate age of the subject at time of entry into the study.

Each participant's online activities were continuously collected throughout the study automatically by the website. Each website activity was time stamped and only activities from consent date to delivery date were included in this analysis. All activities associated with intervention features, rather than data collection activities such as surveys, were considered intervention use in this analysis.

All randomized participants were asked to complete a baseline questionnaire. The questionnaire was available online and via telephone from consent date and until the participant was greater than 28 weeks into pregnancy. A survey item that asked about home Internet use was included in this research.

Conceptualizing Measures of Engagement

Use of the following six intervention features were used to characterize engagement: health-related information (articles and FAQs), blogs, local resources, diet goal-setting tools, physical activity goal-setting tools, and a weight-gain tracker. Features were categorized based on expected use. Consistent use was expected for log-ins and entry of weights into the weight gain tracker. We expected women to track their weight in 30-day intervals but, to allow for difference in timing of doctor’s visits, we created 45-day intervals from time of enrollment to delivery. If a woman entered a weight during each of the 45-day intervals that she completed, she was categorized as a “consistent tracker”. If during at least of half of the intervals a woman entered a weight, she was categorized as an “almost consistent tracker”. If a woman had entered at least one weight but not during more than half of her intervals, she was categorized as an “inconsistent tracker”. Finally, if she never entered a weight during pregnancy, she was categorized as a “non-weight tracker”. The same procedure was used to categorize use for log-ins.

For all other features, consistent use was not expected. Use was expected on an “as needed” basis. Therefore, quantity of use defined engagement for the following features: health-related information, blogs, resources, diet goal-setting tools, and physical activity goal-setting tools. A woman’s engagement was categorized into three levels for each of these features: “high” (≥median among users), “low” (<median among users), or “never” (0).

Demographic Subgroups

Since sociodemographic characteristics are correlated and most sociodemographic characteristics are measured categorically, several recent studies have employed latent class or subgroup analysis to group women with similar characteristics together [15,17,18,24]. Demographic/BMI subgroups were created in the analysis sample (n=1014) based on the following variables: race (white, black, or other), ethnicity (Hispanic or non-Hispanic), low-income status (<185% poverty line or ≥185% poverty line), BMI category (normal (BMI 18-<25), overweight (BMI 25-<30), or obese class 1 (BMI 30-<35), and age category (18-<25 years, 25-<30 years, or 30-35 years).

Home Internet Use

The data for home Internet usage came from the baseline questionnaire survey item: “How often do you access the Internet from your home?” and had the following response categories: never; less than once a week; a few times a week; most days of the week; every day (859/1014, 84.71% of analytic sample). For the purposes of this analysis, we used the following categorizations: “never/occasionally” (never to a few times a week), “most days of the week”, and “every day”.

Analysis

Engagement Patterns

Latent class analysis (LCA) was used to identify patterns of feature use as a measure of overall intervention engagement [25]. Often the latent class variable is used to organize multiple dimensions of behavior, such that individuals in each latent class share common behavior patterns. In our case, this analysis was used to group individuals based on their similar patterns feature use of the intervention website.

LCA models are fit in a series of steps starting with a one-class model; the number of classes is subsequently increased until there is no further improvement in the model. Model selection in LCA involved both absolute fit of a particular model and relative fit of two or more competing models. A common measure of absolute model fit in categorical models is the G^2 likelihood-ratio chi-square statistic, which in our case tests the null hypothesis that the specified LCA model fits the data [26]. Relative fit of models with different numbers of latent classes (eg, 4 vs 5 classes) was analyzed using a series of standard fit indices, including the Bayesian information criterion (BIC [27]) and Akaike information criteria (AIC [28]), with a lower value suggest a more optimal balance between model fit and parsimony. All analyses were conducted using a SAS procedure, PROC LCA [16].

Demographic/Body Mass Index Subgroups

LCA was used to identify demographic/BMI subgroups. Given the strong correlation between demographic and BMI characteristics in this sample, we decided to take a person-centered approach to categorizing women. To do this, we used LCA to identify subgroups within the population based on race, ethnicity, income, BMI, and age. The same LCA model selection criteria were used as with the engagement patterns outlined above.

Association Between Demographic/Body Mass Index Subgroups and Engagement

Chi-square analysis was used to first examine the relationship between individual feature use and demographic subgroup. Next, chi-square analysis was used to examine the relationship between demographic subgroups and patterns of engagement.

The data analysis for this paper was generated using SAS software, version 9.3.

Results

Characterize How Pregnant Women Engaged With Online Intervention Features

The first objective of this study was to capture multiple measures of *how* women used the intervention website. Most women logged into the website during pregnancy (87.97%, 892/1014) and engaged with the intervention features. As described earlier, consistency of use or quantity of use was used to characterize dose for each intervention feature. Of the intervention features, the weight tracker was most commonly used with 25.05% (254/1014) of women who consistently used, 28.99% (294/1014) almost consistently used, 19.03% (193/1014) inconsistently used, and 26.04% (264/1014) never used (Table 2). Health-related information and blogs were engaged by over half of the sample, while the diet and physical activity goal-setting tools were utilized by only a third of the sample.

When all intervention features were considered together, six patterns of engagement emerged from the latent class analysis, as shown in the column headings in Table 3. The first class was characterized by high and consistent usage of all features and is labeled “super-users” (13.02%, 132/1014). “Medium-users” (14.00%, 142/1014) were characterized by almost consistent weight-tracker use and high use of both health-related information and blogs. The next three classes were characterized in the latent class analysis solely based on weight-tracker use: “consistent weight-tracker users” (14.99%, 152/1014); “almost consistent weight-tracker users” (21.99%, 223/1014), and “inconsistent weight-tracker users” (15.98%, 162/1014). The final class, “non-users” (20.02%, 203/1014) were categorized by never engaging with the intervention features.

Table 2. Proportion of total sample (n=1014) that used website features.

Feature use	Analysis sample n (%)
Feature categorized by consistency	
Log-in	
Consistent	332 (32.74)
Almost consistent	342 (33.73)
Inconsistent	214 (21.11)
Never used	126 (12.42)
Weight-tracker	
Consistent	252 (24.85)
Almost consistent	298 (29.39)
Inconsistent	196 (19.33)
Never used	268 (26.43)
Feature categorized by quantity	
Health-related info	
High	270 (26.63)
Low	229 (22.58)
None	515 (50.79)
Blogs	
High	277 (27.32)
Low	272 (26.82)
None	465 (45.86)
Resources	
High	207 (20.41)
Low	175 (17.26)
None	632 (62.33)
Physical activity goal setting	
High	176 (17.36)
Low	139 (13.71)
None	699 (68.93)
Diet goal setting	
High	182 (17.95)
Low	142 (13.81)
None	690 (68.04)

Table 3. Patterns of intervention feature use identified from latent class analysis probabilities.

Feature use	Super-users (13.02%, 132/1014)	Medium-users (14.00%, 142/1014)	Weight-consistent (14.99%, 152/1014)	Weight-almost consistent (21.99%, 223/1014)	Weight-inconsistent (15.98%, 162/1014)	Non-users (20.02%, 203/1014)
Log-in						
Consistent	0.93	0.12	0.97	0.14	0.04	0.00
Almost consistent	0.07	0.86	0.03	0.84	0.08	0.02
Inconsistent	0.00	0.02	0.00	0.01	0.89	0.34
Never	0.00	0.00	0.00	0.00	0.00	0.64
Weight-tracker						
Consistent	0.88	0.00	0.83	0.02	0.00	0.00
Almost consistent	0.11	0.85	0.13	0.63	0.00	0.00
Inconsistent	0.01	0.13	0.01	0.28	0.71	0.00
Never	0.00	0.02	0.03	0.08	0.29	1.00
Physical activity goal setting						
High	0.53	0.45	0.18	0.02	0.04	0.00
Low	0.17	0.21	0.14	0.14	0.20	0.00
None	0.30	0.33	0.67	0.84	0.76	1.00
Diet goal setting						
High	0.50	0.42	0.17	0.05	0.10	0.00
Low	0.16	0.20	0.18	0.15	0.20	0.00
None	0.34	0.38	0.65	0.80	0.70	1.00
Health-related info						
High	0.84	0.65	0.23	0.08	0.06	0.00
Low	0.16	0.27	0.42	0.31	0.20	0.01
None	0.00	0.09	0.35	0.61	0.73	0.99
Blogs						
High	0.91	0.59	0.24	0.09	0.08	0.00
Low	0.07	0.31	0.44	0.43	0.33	0.01
None	0.01	0.09	0.32	0.48	0.59	1.00
Local resources						
High	0.89	0.47	0.06	0.01	0.05	0.00
Low	0.11	0.27	0.33	0.20	0.16	0.00
None	0.00	0.25	0.61	0.79	0.80	1.00

Identify Demographic/Body Mass Index Subgroups of Pregnant Women

The second objective was to create a holistic measure for demographic/BMI characteristics through latent class analysis. From this analysis, four demographic/BMI subgroups emerged, characterized primarily by race, ethnicity, and income (Table 4). The largest subgroup comprised mainly white, non-Hispanic, not low-income, normal weight, and 30 years or older women (54.93% of the sample, 557/1014), which for brevity has been labeled the “white” subgroup. This subgroup was the only subgroup with a high probability of being higher income and older. There were two subgroups that had a high probability of being black. The first comprised black women who were predominantly non-Hispanic, low-income, normal weight, and 18 to 25 years old (20.51%, 208/1014) and is termed “black, young”. The second subgroup comprised black women and was also predominantly non-Hispanic and low-income, but differed by BMI. In this subgroup, women were more likely to be overweight or obese BMI (13.02%, 132/1014). This subgroup was labeled “black, heavier” to denote the distinction between the two subgroups that had a high likelihood of being black. The final group comprised predominantly Hispanic women who were also likely to be low-income, normal, or overweight BMI, and 18 to 25 years old (11.54%, 117/1014). It was the only subgroup that emerged with a greater than 50% probability of Hispanic ethnicity and as such it is labeled “Hispanic”.

Table 4. Demographic/body mass index (BMI) subgroups identified from latent class analysis probabilities.

	Demographic/BMI Subgroup			
	Black, young (20.51%, 208/1014)	Black, heavier (13.02%, 132/1014)	Hispanic (11.54%, 117/1014)	White (54.93%, 557/1014)
Race				
White	0.29	0.42	0.02	0.94
Black	0.71	0.56	0.01	0.00
Other	0.00	0.01	0.96	0.06
Hispanic				
Yes	0.06	0.04	0.84	0.02
No	0.94	0.96	0.16	0.98
Low-income				
Yes	0.89	0.67	0.78	0.11
No	0.11	0.33	0.22	0.89
Body mass index category				
Normal	0.67	0.12	0.41	0.60
Overweight	0.20	0.45	0.41	0.27
Obese	0.13	0.43	0.18	0.12
Age category				
18 - <25	0.82	0.23	0.55	0.07
25 - <30	0.14	0.49	0.31	0.34
≥30	0.03	0.28	0.15	0.59

Examine Differences in Use of Intervention Features Across Demographic/Body Mass Index Subgroups

In the final objective, we examined *who*, based on the demographic/BMI subgroups, engaged with the online intervention. Use of all intervention features was significantly different across demographic/BMI subgroup (Table 5).

The weight-tracker was used by more than half of each of the demographic subgroups. The predominantly white subgroup had the smallest proportion of women who never used it (12.6%, 70/557) and the highest proportion of women who used it consistently (35.7%, 199/557). The minority subgroups also used the weight-tracker but to a lesser degree.

When comparing use of the intervention features together with the LCA patterns, there were pronounced differences across demographic/BMI subgroups (Figure 1). Super-users were more likely to be in the white subgroup compared to other subgroups (20% vs 3%, 9%, and 8%), while non-users were more likely to be in the minority subgroups compared to other subgroups (36%, 36%, and 34% vs 8%). However, around a third of individuals in the minority subgroups were consistently or almost consistently engaging with the weight-tracker (black, young women, 32.2%, 67/208; black, heavier women, 37.9%, 50/132; Hispanic women, 27.4%, 32/117).

While home Internet use behaviors varied across demographic/BMI subgroups (Figure 2), at least 75% of each subgroup used the Internet every day or most days of the week. However, only 26 out of 528 (4.9%) of women in the white subgroup rarely or never used the Internet, while 35 out of 147 (23.8%) of women in the black, young subgroup; 15 out of 98 (15%) of women in the black, heavier subgroup; and 26 out of 86 (25%) of women in the Hispanic subgroup rarely or never used the Internet at home.

Table 5. Intervention feature use by demographic/body mass index (BMI) subgroups.

Feature use		Demographic/BMI Subgroup				*P* value[a]
		Black, young, n=208	Black, heavier, n=132	Hispanic, n=117	White, n=557	
		n (%)	n (%)	n (%)	n (%)	
Log-in						<.01
	Consistent	33 (15.9)	28 (21.2)	28 (23.9)	243 (43.6)	
	Almost consistent	61 (29.3)	39 (29.6)	33 (28.2)	209 (37.5)	
	Inconsistent	64 (30.8)	35 (26.5)	31 (26.5)	84 (15.1)	
	Never used	50 (24.0)	30 (22.7)	25 (21.4)	21 (3.8)	
Weight-tracker						<.01
	Consistent	17 (8.2)	16 (12.1)	20 (17.1)	199 (35.7)	
	Almost consistent.	40 (19.2)	30 (22.7)	27 (23.1)	201 (36.1)	
	Inconsistent	58 (27.9)	28 (21.2)	23 (19.7)	87 (15.6)	
	Never used	93 (44.7)	58 (43.9)	47 (40.2)	70 (12.6)	
Health-related info						<.01
	High	22 (10.6)	19 (14.4)	17 (14.5)	212 (38.1)	
	Low	30 (14.4)	25 (18.9)	25 (21.4)	149 (26.8)	
	None	156 (75.0)	88 (66.7)	75 (64.1)	196 (35.2)	
Blogs						<.01
	High	31 (14.9)	25 (18.9)	22 (18.8)	199 (35.7)	
	Low	44 (21.2)	35 (26.5)	23 (19.7)	170 (30.5)	
	None	133 (63.9)	72 (54.6)	72 (61.5)	188 (33.8)	
Resources						<.01
	High	13 (6.3)	22 (16.7)	12 (10.3)	160 (28.7)	
	Low	26 (12.5)	22 (16.7)	18 (15.4)	109 (19.6)	
	None	169 (81.3)	88 (66.7)	87 (74.4)	288 (51.7)	
Physical activity goal setting						<.01
	High	21 (10.1)	20 (15.2)	11 (9.4)	124 (22.3)	
	Low	21 (10.1)	19 (14.4)	15 (12.8)	84 (15.1)	
	None	166 (79.8)	93 (70.5)	91 (77.8)	349 (62.7)	
Diet goal setting						.009
	High	26 (12.5)	21 (15.9)	17 (14.5)	118 (21.2)	
	Low	23 (11.1)	13 (9.9)	19 (16.2)	87 (15.6)	
	None	159 (76.4)	98 (74.2)	81 (69.2)	352 (63.2)	

[a] Chi-square analysis, *P* value comparing analysis sample and those not included (n=112) from the intervention sample.

Figure 1. Associations between patterns of online engagement and demographic/body mass index (BMI) subgroups (n=1014).

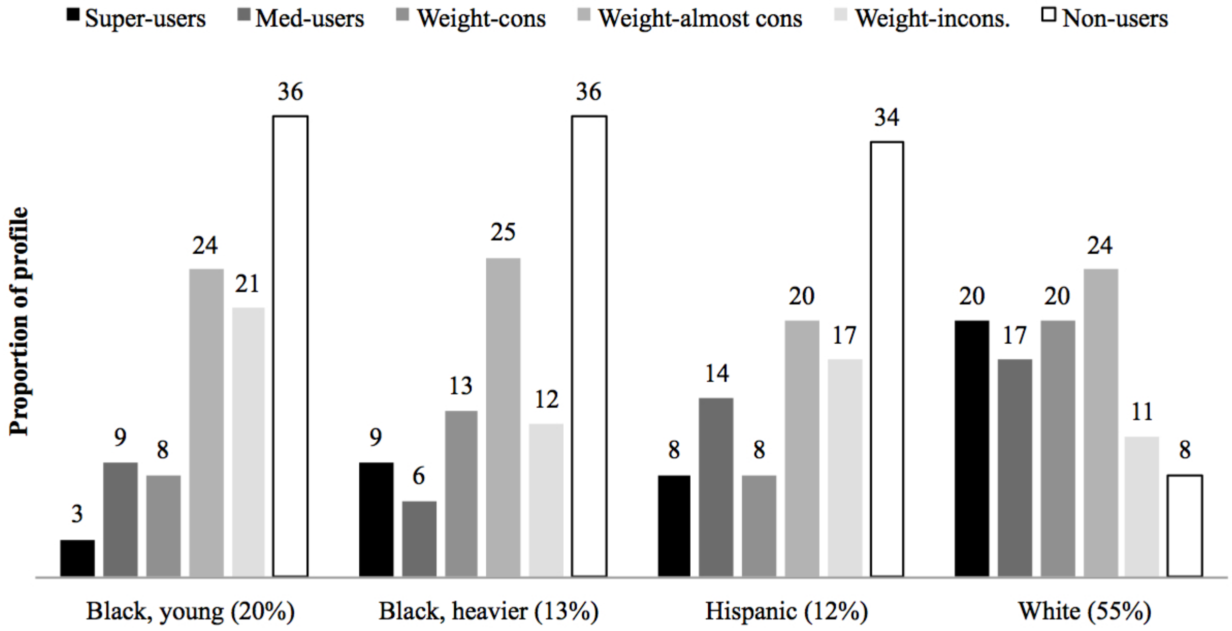
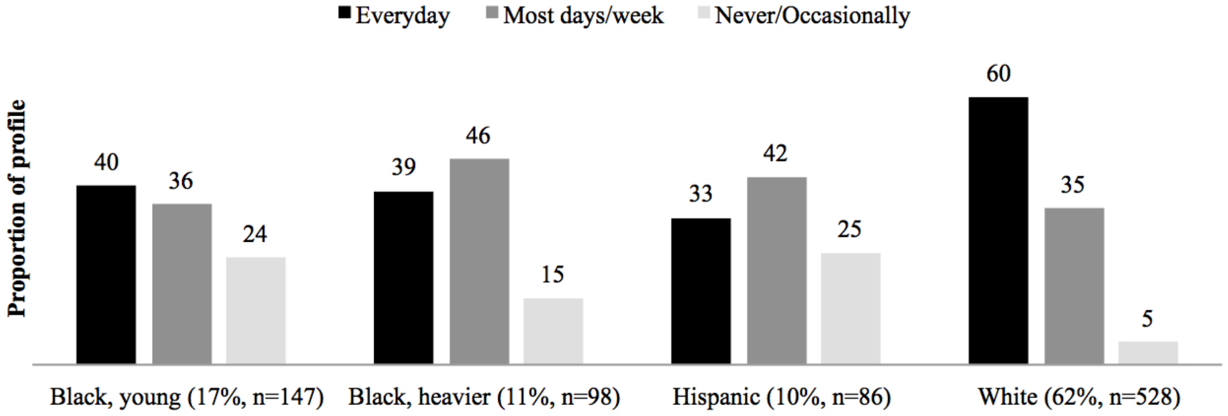


Figure 2. Frequency of home Internet use by demographic/body mass index subgroup (84.71%, 859/1014 women in the analysis sample completed the survey question regarding home Internet use).



Discussion

Principal Findings

Summary

This study examined a variety of engagement measures, based on both how participants used intervention features on the website and expected use of some features, and created data-driven measures of engagement. In addition, subgroups were identified, acknowledging that demographic and weight status characteristics are often associated with each other, allowing for a more person-centered approach. The contributions of this study by aim are outlined in the following sections.

Characterizing and Measuring Engagement

Two methodological contributions were made by the analysis of the online engagement in this study to prevent excessive weight gain in pregnancy. First, consistency of use over time, particularly for weight gain self-monitoring, was expected to be related to appropriate weight gain over time [29-31]. A simple count of 10 weights tracked without taking time into account could mean a woman entered all 10 weights in her first 30 days of the study and discontinued weight monitoring after that. Use of the weight gain tracker was expected to align with each prenatal visit, which typically occurs once a month during the first and second trimesters and bi-weekly in the third trimester. One weight entry per 45-day interval was used to define consistent use across time to allow for variability when participants' doctor's appointments could fall. Characterizing engagement by consistency of use is a unique contribution to capturing online engagement.

Second, by examining the patterns of feature use with latent class analysis, we made a novel contribution to how online interventions could measure engagement. Conceptualizing engagement as patterns rather than independent feature use was conceptually relevant in two ways. First, the website was designed to integrate features by related content. For example, if a woman read a blog about breastfeeding, links to articles or FAQs about breastfeeding appeared next to the blog. Second, the website was designed to offer at least a few features that each participant would use and it was expected that some women would engage with everything (super-users) while others might find only one or two features helpful [21]. The use of latent class analysis allowed us to examine the patterns of usage that emerged from actual use. The results of this study are similar to other studies in that a "super-user" group emerged from our latent class analysis [32-34]. These are individuals who engage with all features in high amounts. The findings of this study move beyond just identifying super-users to identify clustering of feature use at varying intensities.

Interestingly, consistency of weight tracking was a defining feature across most of the patterns. The weight tracking tool was one of the most used features of the website and it was used by a variety of demographic/BMI subgroups. Around a third of the minority subgroups were consistently or almost consistently engaging with the weight-tracker (black, young women, 32%; black, heavier women, 38%; Hispanic women, 28%). The implication for future online interventions is that in order to reach a diverse sample, online interventions need to offer a variety of features and need to acknowledge that engagement in the intervention will vary.

An additional consideration for this analysis is that count data, which could be considered continuous, is challenging to use with parametric methods like factor analysis due to the non-normal distribution of the data [35]. This skewness of engagement data is common to online interventions and is typically handled in other studies by counting use of a feature as ever use or never use for both the use of ever and never use [21,32,36-38]. By categorizing the usage of each feature into no use, low use, and high use, we captured some of the spectrum of use for "as needed" features (eg, blogs, articles, resources) and avoided the challenges of dealing with non-normal data with continuous methods.

Subgroups

Utilizing latent class analysis for creating participant subgroups [17] is a relatively novel approach for examining socioeconomic characteristics of individuals. Utilizing a subgroup analysis methodology like latent class analysis for demographic and weight characteristics, which are known to be correlated, allows for a more holistic characterization of individuals. This analysis sought to understand the website use behaviors of the woman considered more comprehensively, taking into account race, ethnicity, income, age, and BMI together.

The findings from this study suggest that minority and low-income women were less engaged with the website compared to white, higher-income women. However, we also found that a significant proportion of minority and low-income women were also engaging with the website, but not as much or as consistently.

Digital Divide

The digital divide, the inequity between groups in access, use, and knowledge of technology, is an important consideration for online interventions seeking to reach a social and racially diverse population. Even though there are no longer significant differences in both smartphone use and Internet use comparing minorities to whites, there are significant differences by income in Internet and smartphone use [39]. From 2000 to 2011, Internet use increased for blacks (35% to 71%) and for those earning less than US$30,000/year (28% to 62%) [39]. Yet Internet use among households that earn more than US$50,000/year is between 90-97%, while current rates among African Americans or those earning less than US$30,000/year is still much lower [39]. The results presented in this study are consistent with these numbers with about a third of the minority subgroups being non-users in the intervention compared to 8% in our white subgroup.

Limitations

Several methodological choices were made that could be seen as limitations. First, for making the demographic/BMI subgroups, we chose to use only variables that were available for all women. This limited the variables for creating the subgroups to screening variables only. Other subgroup analyses have used up to 40 variables to group participants. Had more variables been available, this may have led to more complex subgroups. Second, the digital divide question for the sample comes from an online survey with a back-up telephone survey. While 85% of the sample answered that question, it is likely that these women were more likely to be Internet users. Third, since this study was conducted with pregnant women, its generalizability is limited particularly given their increased likelihood to seek information online compared to the general population [40].

Gestational weight gain outcomes were not examined as part of this analysis though these outcomes will be examined in relation to both treatment assignment and intervention engagement in future analyses. These future analyses will facilitate understanding whether use of particular intervention features accounts for any overall intervention effect on weight outcomes. Across subgroups there were women who never used the online intervention; exploring why some participants never engaged in the intervention is an area of research needed.

Conclusions

Engagement in online interventions is still a concern for reaching the population of most disadvantaged pregnant women. However, in this study with a population-based sample recruited from healthcare practices, a large number of women were reached and even minority and low-income women engaged to varying degrees with the intervention. Acknowledging both the reach of an online intervention and differential engagement in an online intervention are critical to understanding and interpreting the results of current efficacy trials of online interventions and to the design of future online interventions.

Acknowledgments

The authors' contributions were as follows—MMD, MLG, and CMO designed research; MMD and MLG analyzed data; MMD, MLG, and CMO wrote the paper. All authors read and approved the final manuscript. The National Heart, Lung, and Blood Institute and Eunice Kennedy Shriver National Institute of Child Health and Human Development funded this research. (NIH Grant Number: HL096760). Dr Demment is supported by NIH training grant 5T32HD007331 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Brief description and screenshot image of intervention website features.

[PDF File (Adobe PDF File), 672KB - jmir_v16i8e194_app1.pdf]

References

1. Chao D, Farmer D, Sevick M, Espeland M, Vitolins M, Naughton M. The value of session attendance in a weight-loss intervention. Am J Health Behav 2000;24(6):413-421. [doi: 10.5993/AJHB.24.6.2]
2. Jeffery RW, Bjornson-Benson WM, Rosenthal BS, Lindquist RA, Kurth CL, Johnson SL. Correlates of weight loss and its maintenance over two years of follow-up among middle-aged men. Prev Med 1984 Mar;13(2):155-168. [Medline: 6739444]
3. Tinker LF, Rosal MC, Young AF, Perri MG, Patterson RE, Van Horn L, et al. Predictors of dietary change and maintenance in the Women's Health Initiative Dietary Modification Trial. J Am Diet Assoc 2007 Jul;107(7):1155-1166. [doi: 10.1016/j.jada.2007.04.010] [Medline: 17604744]
4. Alexander GL, McClure JB, Calvi JH, Divine GW, Stopponi MA, Rolnick SJ, MENU Choices Team. A randomized clinical trial evaluating online interventions to improve fruit and vegetable consumption. Am J Public Health 2010 Feb;100(2):319-326 [FREE Full text] [doi: 10.2105/AJPH.2008.154468] [Medline: 20019315]
5. 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]
6. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
7. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39. [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
8. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. Web-based smoking-cessation programs: results of a randomized trial. Am J Prev Med 2008 May;34(5):373-381 [FREE Full text] [doi: 10.1016/j.amepre.2007.12.024] [Medline: 18407003]
9. Funk KL, Stevens VJ, Appel LJ, Bauck A, Brantley PJ, Champagne CM, et al. Associations of internet website use with weight change in a long-term weight loss maintenance program. J Med Internet Res 2010;12(3):e29 [FREE Full text] [doi: 10.2196/jmir.1504] [Medline: 20663751]
10. Guidelines CTATPHSOCP, Medicine IO, Council NR, Rasmussen KB, Yaktine AL. Weight gain during pregnancy: reexamining the guidelines. Washington, DC: National Academies Press; 2009.
11. Rooney BL, Schauberger CW. Excess pregnancy weight gain and long-term obesity: one decade later. Obstet Gynecol 2002 Aug;100(2):245-252. [Medline: 12151145]
12. Linné Y, Dye L, Barkeling B, Rössner S. Long-term weight development in women: a 15-year follow-up of the effects of pregnancy. Obes Res 2004 Jul;12(7):1166-1178. [doi: 10.1038/oby.2004.146] [Medline: 15292482]
13. 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):130-136. [Medline: 16542624]
14. Linde JA, Jeffery RW, French SA, Pronk NP, Boyle RG. Self-weighing in weight gain prevention and weight loss trials. Ann Behav Med 2005 Dec;30(3):210-216. [doi: 10.1207/s15324796abm3003_5] [Medline: 16336072]
15. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci 2013 Apr;14(2):157-168 [FREE Full text] [doi: 10.1007/s11121-011-0201-1] [Medline: 21318625]
16. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. Struct Equ Modeling 2007;14(4):671-694 [FREE Full text] [Medline: 19953201]
17. Scharoun-Lee M, Gordon-Larsen P, Adair LS, Popkin BM, Kaufman JS, Suchindran CM. Intergenerational profiles of socioeconomic (dis)advantage and obesity during the transition to adulthood. Demography 2011 May;48(2):625-651 [FREE Full text] [doi: 10.1007/s13524-011-0024-5] [Medline: 21491185]

18. Brody GH, Yu T, Chen YF, Kogan SM, Evans GW, Beach SR, et al. Cumulative socioeconomic status risk, allostatic load, and adjustment: a prospective latent profile analysis with contextual and genetic protective factors. Dev Psychol 2013 May;49(5):913-927 [FREE Full text] [doi: 10.1037/a0028847] [Medline: 22709130]
19. Fishbein M, Yzer MC. Using theory to design effective health behavior interventions. Communication Theory 2003;13(2):164-183. [doi: 10.1111/j.1468-2885.2003.tb00287]
20. Fogg B. Persuasive technology. Ubiquity 2002 Dec 01;2002(December):2. [doi: 10.1145/764008.763957]
21. Graham ML, Uesugi KP, Niederdeppe J, Gay GK, Olson CM. The theory, development, and implementation of an online intervention to prevent excessive gestational weight gain: eMoms Roc. Telemedicine Journal and E-Health 2014 (forthcoming).
22. Lytle L, Svetkey LP, Patrick K, Belle S, Fernandez I, Jakicic J, et al. The EARLY trials: a consortium of studies targeting weight control in young adults. Translational Behavioral Medicine 2014 (forthcoming).
23. e-Moms of Rochester Website. 2011. e-Moms of Rochester Healthy Pregnancy Study URL: http://www.emomsroc.org/ [accessed 2014-04-21] [WebCite Cache ID 6P0akSs9n]
24. Dewilde C. The multidimensional measurement of poverty in Belgium and Britain: a categorical approach. Social Indicators Research 2004 Sep;68(3):331-369. [doi: 10.1023/B:SOCI.0000033578.81639.89]
25. Nagin D. Group-based modeling of development. Cambridge, Mass: Harvard University Press; 2005.
26. Agresti A. Categorical Data Analysis (Wiley Series in Probability and Statistics). Hoboken, New Jersey: Wiley; 2012.
27. Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978;6:461-464. [doi: 10.2307/2958889]
28. Akaike H. A new look at the statistical model identification. IEEE Trans. Automat. Contr 1974 Dec;19(6):716-723. [doi: 10.1109/TAC.1974.1100705]
29. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
30. Olson CM, Strawderman MS, Reed RG. Efficacy of an intervention to prevent excessive gestational weight gain. Am J Obstet Gynecol 2004 Aug;191(2):530-536. [doi: 10.1016/j.ajog.2004.01.027] [Medline: 15343232]
31. Olson CM. Achieving a healthy weight gain during pregnancy. Annu Rev Nutr 2008;28:411-423. [doi: 10.1146/annurev.nutr.28.061807.155322] [Medline: 18422452]
32. 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]
33. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. J Med Internet Res 2012;14(3):e66 [FREE Full text] [doi: 10.2196/jmir.1854] [Medline: 22732103]
34. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. J Med Internet Res 2013;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
35. Bartholomew D, Knott M, Moustaki I. Latent Variable Models and Factor Analysis: A Unified Approach (Wiley Series in Probability and Statistics). Hoboken, New Jersey: Wiley; 2011.
36. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. J Med Internet Res 2007;9(2):e11 [FREE Full text] [doi: 10.2196/jmir.9.2.e11] [Medline: 17513282]
37. van Zutphen M, Milder IE, Bemelmans WJ. Usage of an online healthy lifestyle program by pregnant women attending midwifery practices in Amsterdam. Prev Med 2008 Jun;46(6):552-557. [doi: 10.1016/j.ypmed.2008.01.003] [Medline: 18302971]
38. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. J Med Internet Res 2010;12(4):e52 [FREE Full text] [doi: 10.2196/jmir.1430] [Medline: 21087922]
39. Zickuhr K, Smith A. Pew Research Internet Project. 2012 Apr 13. Digital differences URL: http://www.pewinternet.org/2012/04/13/digital-differences/ [accessed 2014-08-13] [WebCite Cache ID 6RnnEyzjF]
40. Larsson M. A descriptive study of the use of the Internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]

Abbreviations

BMI: body mass index
GWG: gestational weight gain
LCA: latent class analysis

Edited by G Eysenbach; submitted 22.04.14; peer-reviewed by V Stevens, J Willcox; comments to author 16.07.14; accepted 22.07.14; published 20.08.14.

Please cite as:
Demment MM, Graham ML, Olson CM
How an Online Intervention to Prevent Excessive Gestational Weight Gain Is Used and by Whom: A Randomized Controlled Process Evaluation
J Med Internet Res 2014;16(8):e194
URL: http://www.jmir.org/2014/8/e194/
doi:10.2196/jmir.3483
PMID:25143156

Original Paper

Mobile Phone Text Messaging Intervention for Cervical Cancer Screening: Changes in Knowledge and Behavior Pre-Post Intervention

Hee Yun Lee[1*], PhD; Joseph S Koopmeiners[2*], PhD; Taeho Greg Rhee[3*], AM; Victoria H Raveis[4*], PhD; Jasjit S Ahluwalia[5*], MD, MPH, MS

[1] School of Social Work and University of Minnesota Masonic Cancer Center, College of Education and Human Development, University of Minnesota, Twin Cities, St Paul, MN, United States

[2] Division of Biostatistics, School of Public Health, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[3] Division of Health Policy and Management, School of Public Health, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[4] Psychosocial Research Unit on Health, Aging, and the Community, College of Dentistry, New York University, New York, NY, United States

[5] Center for Health Equity and Department of Medicine, School of Medicine, University of Minnesota, Twin Cities, Minneapolis, MN, United States

[*] all authors contributed equally

Corresponding Author:
Hee Yun Lee, PhD
School of Social Work and University of Minnesota Masonic Cancer Center
College of Education and Human Development
University of Minnesota, Twin Cities
1404 Gortner Ave
St Paul, MN, 55108
United States
Phone: 1 612 624 3689
Fax: 1 612 624 3744
Email: hylee@umn.edu

Abstract

Background: Cervical cancer poses a significant threat to Korean American women, who are reported to have one of the highest cervical cancer mortality rates in the United States. Studies consistently report that Korean American women have the lowest Pap test screening rates across US ethnic groups.

Objective: In response to the need to enhance cervical cancer screening in this vulnerable population, we developed and tested a 7-day mobile phone text message-based cervical cancer Screening (mScreening) intervention designed to promote the receipt of Pap tests by young Korean American women.

Methods: We developed and assessed the acceptability and feasibility of a 1-week mScreening intervention to increase knowledge of cervical cancer screening, intent to receive screening, and the receipt of a Pap test. Fogg's Behavior Model was the conceptual framework that guided the development of the mScreening intervention. A series of focus groups were conducted to inform the development of the intervention. The messages were individually tailored for each participant and delivered to them for a 7-day period at each participant's preferred time. A quasi-experimental research design of 30 Korean American women aged 21 to 29 years was utilized with baseline, post (1 week after the completion of mScreening), and follow-up (3 months after the completion of mScreening) testing.

Results: Findings revealed a significant increase in participants' knowledge of cervical cancer ($P<.001$) and guidelines for cervical cancer screening ($P=.006$). A total of 23% (7/30) (95% CI 9.9-42.3) of the mScreening participants received a Pap test; 83% (25/30) of the participants expressed satisfaction with the intervention and 97% (29/30) reported that they would recommend the program to their friends, indicating excellent acceptability and feasibility of the intervention.

Conclusions: This study provides evidence of the effectiveness and feasibility of the mScreening intervention. Mobile technology is a promising tool to increase both knowledge and receipt of cervical cancer screening. Given the widespread usage of mobile phones among young adults, a mobile phone-based health intervention could be a low-cost and effective method of reaching populations with low cervical cancer screening rates, using individually tailored messages that cover broad content areas and overcome restrictions to place and time of delivery.

(J Med Internet Res 2014;16(8):e196) doi:10.2196/jmir.3576

KEYWORDS

cervical cancer; Pap test; mobile health; text-messaging intervention; health behavior change; Korean American women; health disparity

Introduction

Korean American women have one of the highest cervical cancer mortality rates in the United States. Cervical cancer incidence and mortality rates for Korean American women are roughly twice that of non-Latino white women [1]. While the *Healthy People 2020* initiative states that 93% of women, aged 21-65 years, should have undergone a Pap test within the past 3 years [2], studies consistently report that, among women across US racial/ethnic groups, Korean American women have the lowest Pap test screening rates, ranging from 39% to 64% [3-8]. Given that early detection of cervical malignancies through this routine screening measure has been shown to significantly reduce cervical cancer mortality, Korean American women's low screening rate indicates that efforts to increase their screening behavior would be very beneficial [1,9].

A variety of structural and cultural factors act as barriers to screening for Korean American women. Structural obstacles include health access due to inadequate health insurance [10-12], expense [1,13], time constraints [10,13], and language limitations [1,10,12-15]. Cultural barriers to cervical cancer screening encompass lack of knowledge regarding cervical cancer and cervical cancer screening [1,10,12,13,16], a wrongly held belief that screening is unnecessary in the absence of symptoms or at young ages [1,10,12,13,16,17], cultural modesty or embarrassment [10,13,16], lack of culturally appropriate health care providers [12,13], and fear of receiving negative screening results [10,16].

A limited number of interventions to address barriers and promote cervical cancer screening among Korean American women have been designed or implemented. These efforts have focused on peer-led workshops [18,19], dissemination of videos [18], and distribution of cancer education print materials [20,21]. There are a number of reasons why these approaches have been only partially effective in promoting cervical cancer screening in this population. Korean American women are a particularly hard-to-reach population [19,20]. Although earlier interventions have specifically targeted structural barriers to cancer screening (eg, providing low-cost or free Pap tests or in-language services), prominent cultural obstacles such as cultural modesty or misconceptions about screening were not addressed [19]. These previous intervention strategies also did not tailor the intervention to strategically target specific individual concerns about screening, despite evidence that there are multiple cultural reasons for Korean American women's reservations about cervical cancer screening [20]. The restricted scope and lack of tailoring in these previous interventions may have contributed to their limited impact. Personalized interventions may be necessary to motivate a change in screening behavior with this difficult to reach population.

To address the multiple limitations that were present in prior interventions, we developed and tested a mobile phone text message-based cervical cancer Screening (mScreening) intervention that utilizes mobile health (mHealth) technology. mHealth is defined as the use of mobile and wireless devices as intervention tools to deliver health information or improve health outcomes, particularly using short message service (SMS or text) and/or multimedia message service (MMS, or images or pictures) [22]. mHealth is considered a promising tool for preventive care through promotion of behavioral change. mHealth is taking a primary place in a number of research initiatives related to the promotion of health behavior. For example, mHealth technology was successfully used in weight management [23], smoking cessation [24-28], youth sexual health [29], increased physical activity [30], self-care behaviors [27,28,31], and asthma monitoring and management programs [32].

Our study seeks to harness mobile phone technology to positively influence cancer screening behavior, taking preventive health care approaches to a new level [7,33]. Guided by the Fogg's Behavior Model (FBM) [33], the mScreening intervention consists of three sequential components: (1) identifying barriers, (2) developing motivators, and (3) providing triggers (see Figure 1 for conceptual framework). With the FBM framework, we first identified specific structural and cultural barriers that prevent Korean American women from receiving a Pap test. This information guided our subsequent development and implementation of mobile tools (eg, SMS or MMS) to improve knowledge and provide the motivation for behavioral change. Triggers to act are employed in the form of reminder text messages (eg, "Make an appointment now!") or electronic links (eg, "Click to talk to a Korean health navigator") to prompt Korean American women to take immediate action to obtain a Pap test.

Our study had three primary aims: (1) to examine if the mScreening intervention increased research participants' (a) knowledge of cervical cancer, relevant guidelines of cervical cancer screening, and cervical cancer risk factors, and (b) their intent to undergo screening; (2) to assess if the mScreening intervention contributed to a 20% increase in the receipt of the Pap test over the sample's baseline rate; and (3) to examine the acceptability and satisfaction of the 7-day mScreening intervention program. Our study used evidence-based and theory-driven approaches to develop the mScreening intervention to concentrate on barriers (eg, cultural beliefs, perceived-risk, and limited health literacy) that prior work had not addressed, so as to increase Korean American women's adherence to cervical cancer screening guidelines.

Figure 1. Conceptual framework of the study.

Methods

Research Design and Data Collection

A quasi-experimental design was used with baseline, post-test (1 week after the completion of mScreening), and follow-up (3 months after the completion of mScreening) assessments. The eligibility criteria for our study included: Korean American women aged 21 to 29 years with no prior receipt of a Pap test in Minnesota. Participants also needed to have up-to-date health insurance, mobile phone access (either a regular or smartphone), and be familiar with or willing to learn basic elements of text-based information communication technology. We used a multi-pronged recruitment strategy. A variety of outreach flyers and brochures were posted and handed out in churches, clinics, social service agencies, and ethnic markets that serve the Korean American population, as well as advertised through public social media that is tailored towards the Korean American population. We fully explained the purpose of our study, eligibility criteria, confidentiality, and the voluntary nature of study participation to every participant both in oral and written formats before the study began. The University of Minnesota Institutional Review Board approved the research procedure for this study. A total of 30 participants were enrolled who met the inclusion criteria, received the intervention, and completed the data collection protocol. We had complete retention of participants in our study.

Intervention

We used a community-based participatory research approach to develop the content of the 7-day text message-based mScreening intervention and participant recruitment strategies. We formed a community advisory board, which consisted of Korean American community leaders, health care professionals who work in the Korean American community, and directors of social service agencies. Additionally, we conducted a series of focus groups with 13 young Korean American women. After usability testing for the mScreening intervention with 8 Korean American women, we recruited 34 Korean American women and delivered the mScreening intervention to the 30 women who qualified as study participants.

As per the FBM [33], the mScreening intervention identifies individual barriers to screening, develops motivators specific to these barriers, and provides a trigger to the desired health behavior action (ie, receipt of a Pap test) (see Figure 1). The informational/educational content of the mScreening intervention covered the following topics: (1) introductory information on the cervix and cervical cancer, including statistical facts of cervical cancer incidence and mortality, and screening rates of cervical cancer among Korean American women compared to other racial/ethnic groups in the United States, (2) introduction of the Pap test as a preventive mode for cervical cancer, (3) information on health care accessibility, (4) introduction of cultural barriers, (5) availability of local clinics and cost of Pap test, (6) testimony of a Korean American woman who had gone through the Pap test experience, and (7) testimony of a Korean American cervical cancer survivor who found cervical cancer at later stage and had no previous receipt of a Pap test. The mScreening intervention was delivered to each participant over 20-30 minutes each day for a 7-day period at each participant's preferred time.

The mScreening intervention contained a high level of interactive features, such as quizzes and questions, and also allowed participants to engage in conversation, with approximately one-third of messages asking for responses. Messages were individually tailored for each participant. For example, based on the baseline interview of each participant, we identified strengths and weaknesses of individual participants. If a participant had weaknesses on culture-based health beliefs on cervical cancer screening at baseline, we provided ad hoc messages that were designed to reduce cultural barriers (eg, "We understand it is a bit embarrassing to get it done. But do it for you! Your happy cervix will appreciate it!") in addition to regular messages sent each day. This tailoring was enabled by a database of over 50 ad hoc messages that were designed to reduce culture-based beliefs and attitudes on cervical cancer screening.

Measures

Overview

We collected information from participants through face-to-face interviews using a structured questionnaire administered at three time points: at the study enrollment (baseline), 1 week after completing the mScreening intervention (post-test), and 3 months after completing the mScreening intervention (follow-up).

Outcome Measures

Our primary outcomes of interest included knowledge, attitudes, and beliefs about cervical cancer and cervical cancer screening, intent to undergo screening as measured with the trans-theoretical model (eg, stages of change consisting of precontemplation, contemplation, preparation for action, action, and maintenance) [34,35], and the receipt of the Pap test.

Baseline Measures

We collected information on sociodemographic characteristics (eg, age, marital status, education, employment status, income, health status, and religion), health-related information (eg, history of cancer in the family, health insurance, usual source of care, and number of doctor visits), level of acculturation (years in the United States and English language proficiency), and administered scales for health knowledge and beliefs about cervical cancer, cervical cancer screening, and intent to undergo a Pap test as part of cervical cancer screening. Champion's revised health belief model scale was used for measuring health beliefs on cervical cancer and cervical cancer screening [36]. This scale was previously used in Korean American women in Chicago and was demonstrated to have appropriate reliability and validity [1]. We adapted and used Taylor and colleague's 15-item scale for knowledge about cervical cancer and cervical cancer screening [37].

Post-Test and Follow-Up Measures

One week after the completion of the mScreening intervention (post-test interview), we asked participants about their experience with the intervention and general satisfaction to inform quality assurance and improvement. We also repeated the items from the baseline interview about knowledge and beliefs about cervical cancer and cervical cancer screening, and assessed their intent post-test to undergo screening and if they received a Pap test after the intervention. In the 3-month follow-up interview, we asked participants if they had received a Pap test in the prior 3 months, and for those who did not receive a Pap test, we asked their reasons for not undergoing the test.

Data Analyses

Differences in constructs relating to knowledge, attitudes, and beliefs about cervical cancer and cervical cancer screening from baseline to 1 week post-test were summarized by means and standard deviations, and tested using the paired *t* test. Two approaches for comparing the intent to receive the Pap test pre- and post-mScreening were considered. First, we translated the change in a subject's intent to receive the Pap test to a numerical scale with subjects receiving a "1" if their intent to receive the Pap test increased from pre- to post-mScreening, a "0" if their intent to receive the Pap test stayed the same, and a "−1" if their intent to receive the Pap test decreased. We then summarized subjects' change in their intent to receive the Pap test by the mean and standard deviation and compared the mean change to zero using the one-sample *t* test to determine if, on average, subjects' intent to receive the Pap test increased from pre- to post-mScreening. In addition, we dichotomized intent to receive the Pap test into "intent to receive within a year" and "no intent to receive within a year". The percent of subjects providing each response before and after mScreening were compared using McNemar's test for paired binary data. We estimated the rate of cervical cancer screening post-intervention and at the 3-month follow-up using the sample proportion and 95% confidence intervals were estimated using the exact method. Measures of acceptability and satisfaction were summarized using counts and sample proportions. Given the preliminary nature of the study and relatively small sample size, multivariate analyses were not conducted and only univariate and bivariate results are reported.

Results

Demographic Characteristics of the Sample

Table 1 presents the study participants' baseline demographic information (N=30). Participants ranged in age from 21-29 years, and 27 out of 30 participants (90%) reported living in the United States for less than 9 years, with 37% (11/30) living in the United States for 3 years or less. Of the 30 participants, 28 participants (93%) reported that they had health insurance; 15 participants (50%) reported having a primary hospital; and 6 participants (20%) reported having a primary physician. Finally, 19 participants (63%) reported a family history of cancer.

Table 1. Baseline sociodemographic characteristics of the sample (N=30).

Variable	Categories	n (%)
Age, years		
	21-22	11 (37)
	23-25	13 (43)
	26-29	6 (20)
Marital status		
	Married	1 (3)
	Other	29 (97)
Years in United States		
	3 or less	11 (37)
	4-8	16 (53)
	9 or more	3 (10)
Employment		
	Yes	11 (37)
	No	19 (63)
Education		
	Undergraduates	19 (63)
	Graduated from college/university	9 (30)
	Graduated from graduate school	2 (7)
Monthly income (US$)		
	under $499	14 (48)
	$500-$1499	10 (34)
	$1500 or more	5 (17)
Health status		
	Poor or fair	19 (63)
	Good	6 (20)
	Very good or excellent	5 (17)
Living arrangement		
	Live alone	12 (40)
	Live with spouse only	1 (3)
	Live with others	17 (57)
	Other	0 (0)
Currently living in		
	Rented house or condominium	19 (63)
	Government-subsidized senior citizen apartment	4 (13)
	Unsubsidized apartment	3 (10)
	Rented room in other’s home	3 (10)
	Other	1 (3)
Self-rated financial status		

Variable	Categories	n (%)
	Very bad	1 (3)
	Bad	3 (10)
	Fair	20 (67)
	Good	4 (13)
	Very good	2 (7)
How often do you work out per week?		
	Not at all	8 (27)
	Once	2 (7)
	Twice	11 (37)
	Three times	4 (13)
	Four times or more	5 (17)
Smoking status		
	Not at all	28 (93)
	Some days	1 (3)
	Every day	1 (3)
Alcohol consumption		
	Not at all	13 (43)
	Some days	17 (57)
	Every day	0 (0)
Health insurance		
	Yes	28 (93)
	No	2 (7)
Primary hospital		
	Yes	15 (50)
	No	15 (50)
Primary physician		
	Yes	6 (20)
	No	24 (80)
Family cancer history		
	Yes	19 (63)
	No	11 (37)

Changes in Knowledge, Attitudes, and Beliefs

Table 2 presents changes in measures of knowledge, attitudes, and beliefs about the Pap test, comparing baseline to the post-test, completed 1 week after the mScreening intervention. Significant differences were observed for all constructs and these differences remained significant after a Bonferonni multiple comparison adjustment (*P*<.01). Significant improvements were observed for general knowledge about cervical cancer (*P*<.001), knowledge about the Pap test (*P*<.001), beliefs about and attitudes toward the Pap test (*P*=.006), and knowledge about risk factors of cervical cancer and its screening (*P*<.001). We also observed a significant reduction in socio-cultural barriers to cervical cancer screening (*P*=.001).

Table 2. Changes in knowledge, attitudes, and beliefs about cervical cancer and the Pap test (N=30).

Construct	Baseline pretest mean (SD)	1 week posttest mean (SD)	Mean difference	*P* value
General knowledge about cervical cancer	0.33 (0.36)	0.92 (0.17)	0.59	<.001
Knowledge about Pap test	2.85 (0.87)	3.56 (0.52)	0.71	<.001
Beliefs of and attitude toward Pap test	2.83 (0.42)	3.14 (0.52)	0.31	.006
Knowledge about risk factors of cervical cancer and its screening	2.11 (0.80)	2.87 (0.57)	0.76	<.001
Socio-cultural barriers to cervical cancer screening	2.33 (0.51)	1.98 (0.53)	−0.35	.001

Intent and Receipt of Pap Test

Table 3 presents the study participants' intent to receive the Pap test before and after the mScreening program. We observed an increase in participants' intent to receive the Pap test (mean 0.23, 95% CI −0.04 to 0.51) and the percent of participants indicating an intent to receive the Pap test within 1 year (from 63% to 87%) but these differences were not statistically significant (*P*=.090 and *P*=.070, respectively). In addition, one study participant (3%, 1/30) reported receiving the Pap test within 1 week after completing the mScreening program and 6 additional participants (20%, 6/30) reported receiving the Pap test by the 3-month follow-up visit, which represents a 23% (7/30) improvement in the proportion of participants receiving the Pap test (95% CI 10% to 42%) compared to baseline.

Table 3. Intention to receive the Pap test before and after the mScreening intervention (N=30).

Baseline pretest	Intent to receive a Pap test at 1 week post-test, n (%)					
	No plans within 1 year	Within 1 year	Within 3 months	Within 1 month	Have received Pap test	Total
No plans within 1 year	2 (6.7)	4 (13.3)	2 (6.7)	2 (6.7)	1 (3.3)	11 (36.7)
Within 1 year	2 (6.7)	9 (30.0)	1 (3.3)	1 (3.3)	0 (0.0)	13 (43.3)
Within 3 months	0 (0.0)	2 (6.7)	0 (0.0)	1 (3.3)	0 (0.0)	3 (10.0)
Within 1 month	0 (0.0)	1 (3.3)	0 (0.0)	2 (6.7)	0 (0.0)	3 (10.0)
Have received Pap test	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	4 (13.4)	16 (53.3)	3 (10.0)	6 (20.0)	1 (3.3)	30 (100)

Acceptability and Feasibility of the mScreening Intervention

Table 4 presents participant responses to questions relating to the acceptability and satisfaction of the 7-day mScreening intervention. At the 1-week post-test visit, 25 of 30 participants (83%) reported being either satisfied or very satisfied with the mScreening program and 29 participants (97%) reported that they would recommend mScreening to their friends.

Table 4. Acceptability and satisfaction of the mScreening intervention at 1-week post-test interview (N=30).

Question	Answer	n (%)
Please rate your satisfaction level with mScreening program		
	Very satisfied	8 (27)
	Satisfied	17 (57)
	Neutral	5 (17)
	Dissatisfied	0 (0)
	Very dissatisfied	0 (0)
Would you like to recommend mScreening program to your friend?		
	Yes	29 (97)
	No	1 (3)

Discussion

Principal Findings

We observed a significant increase in the study sample's general knowledge about cervical cancer and the Pap test, beliefs and attitudes toward the Pap test, and knowledge about risk factors of cervical cancer and its screening guidelines (*P*=.006), after completing the mScreening intervention. We also found a significant decrease in perceived socio-cultural barriers to cervical cancer screening (*P*=.005). By the 3-month follow-up, 7 out of 30 participants (23%) reported having received the Pap test. As only women that had not previously received the Pap test were recruited for this study, this finding indicates that participation in the mScreening intervention led to more than a 20% increase in the receipt of the Pap test in this sample. This is a notable achievement as we targeted a group that had not previously been motivated to engage in this recommended health promotion activity. These findings are in line with a previous study, where an SMS reminder system resulted in a significant increase in the practice of breast self-examination [38].

Positive results were also obtained regarding the acceptability and satisfaction of the 7-day mScreening intervention program. At the 1-week post-test visit, 83% (25/30) of participants reported that they were either satisfied or very satisfied with the mScreening program, and 97% (29/30) of participants stated that they would recommend the mScreening program to their friends. Overall, the majority of participants expressed a high degree of acceptability and satisfaction with the mScreening intervention program.

Our study shows that the FBM [33] was useful in explaining and improving cervical cancer screening behaviors. We identified structural and cultural factors (eg, language limitations and lack of knowledge regarding cervical cancer and cervical cancer screening) as well as individual factors (eg, family history of cancer) that are barriers to cervical cancer screening in Korean American women. The mScreening intervention addressed such factors with the intention of minimizing the barriers and motivating research participants to act upon the preventive health behavior. Triggers (ie, a SMS/MMS reminder) were also provided as part of the mScreening intervention, so that research participants did not only change in terms of their knowledge, attitudes, and/or beliefs about cervical cancer and its screening, but were also actively engaged in the utilization of the Pap test.

Limitations

While the results of this study are very promising, there are some limitations. First, the sample size was relatively small and we used a quasi-experimental study design. Further research is needed to validate the effectiveness of the mScreening intervention with a larger sample of Korean American women using a rigorous research design, such as a randomized controlled trial. Second, our study was not designed to explore what would be the optimum time interval for program delivery. We delivered the mScreening intervention over a 7-day period. The study participants' feedback during the post-test interviews was that a shorter intervention period would have been better. Investigations comparing the 7-day text-message program with a shorter intervention (eg, 5 days or 3 days) may determine if a briefer time may still be sufficient to bring about behavior change. However, we also postulate that an individually tailored intervention with a longer time period may be more effective for those with more barriers. Third, the ideal intervention medium is not yet known. Further research is needed to determine if it is beneficial to individually tailor the length of the intervention and if a more effective medium may result in greater behavior change (such as an interactive smartphone app). Additional investigations should examine the utility of delivering the text-message program via mobile application (mobile app), which utilizes a password function to protect participants' confidentiality and privacy. The efforts to find the best intervention medium (text vs mobile app) and the most appropriate length of intervention (1 week vs fewer days) will create an intervention that is more effective in promoting cervical cancer screening and prevention while at the same time increasing participants' satisfaction.

Conclusions

Our findings revealed that mobile technology is a promising tool to increase both knowledge about cervical cancer and receipt of the Pap test. This study provides evidence for the feasibility, acceptability, and satisfaction of the mScreening intervention. Given the widespread use of mobile phones (98%) and smartphones (83%) among young adults [39], a mobile phone-based health intervention could be a cost-effective method of reaching hard-to-reach populations with tailored, individual messages that cover broad content areas and overcome restrictions to place and time of delivery. Our developed model could be expanded for delivery to different age groups of Korean American women to promote additional types of cancer screening, such as colonoscopy or mammogram. It could also be used with other underserved minority groups. Vietnamese, Hmong, and Laotian American women face similar barriers to cancer screening and report high cervical cancer incidence and mortality [40-42]. It is likely that these populations may also benefit from a similarly tailored intervention approach. Given emerging technological developments, effective interventions that could be adapted to efficiently disseminate culturally appropriate health information and promote positive health behavior changes would broadly impact the social determinants of health disparities in hard-to-reach, vulnerable populations.

Acknowledgments

Funding for this research was provided by a grant from the National Cancer Institute (R21-1R21CA155531-0). Dr Ahluwalia is supported in part by the National Institute for Minority Health Disparities (NIMHD/NIH - 1P60MD003422). The authors would like to extend their gratitude to the Korean American women who participated in this study.

Conflicts of Interest

None declared.

References

1. Lee EE, Fogg L, Menon U. Knowledge and beliefs related to cervical cancer and screening among Korean American women. West J Nurs Res 2008 Dec;30(8):960-974 [FREE Full text] [doi: 10.1177/0193945908319250] [Medline: 18539883]
2. U.S. Department of Health and Human Services. Healthy People 2020 URL: http://www.healthypeople.gov/2020/default.aspx [accessed 2014-07-18] [WebCite Cache ID 6R9OMmF1j]
3. Kagawa-Singer M, Pourat N, Breen N, Coughlin S, Abend McLean T, McNeel TS, et al. Breast and cervical cancer screening rates of subgroups of Asian American women in California. Med Care Res Rev 2007 Dec;64(6):706-730. [doi: 10.1177/1077558707304638] [Medline: 17804823]
4. McCracken M, Olsen M, Chen MS, Jemal A, Thun M, Cokkinides V, et al. Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities. CA Cancer J Clin 2007;57(4):190-205. [Medline: 17626117]
5. Juon HS, Choi S, Klassen A, Roter D. Impact of breast cancer screening intervention on Korean-American women in Maryland. Cancer Detect Prev 2006;30(3):297-305. [doi: 10.1016/j.cdp.2006.03.008] [Medline: 16870356]
6. Chen JY, Diamant AL, Kagawa-Singer M, Pourat N, Wold C. Disaggregating data on Asian and Pacific Islander women to assess cancer screening. Am J Prev Med 2004 Aug;27(2):139-145. [doi: 10.1016/j.amepre.2004.03.013] [Medline: 15261901]
7. Juon HS, Choi Y, Kim MT. Cancer screening behaviors among Korean-American women. Cancer Detect Prev 2000;24(6):589-601. [Medline: 11198273]
8. Maxwell AE, Bastani R, Warda US. Demographic predictors of cancer screening among Filipino and Korean immigrants in the United States. Am J Prev Med 2000 Jan;18(1):62-68. [Medline: 10808984]
9. Phipps E, Cohen MH, Sorn R, Braitman LE. A pilot study of cancer knowledge and screening behaviors of Vietnamese and Cambodian women. Health Care Women Int 1999;20(2):195-207. [doi: 10.1080/073993399245881] [Medline: 10409988]
10. Kim H, Lee KJ, Lee SO, Kim S. Cervical cancer screening in Korean American women: findings from focus group interviews. Taehan Kanho Hakhoe Chi 2004 Jun;34(4):617-624. [Medline: 15502427]
11. Lee HY, Ju E, Vang PD, Lundquist M. Breast and cervical cancer screening disparity among Asian American women: does race/ethnicity matter [corrected]? J Womens Health (Larchmt) 2010 Oct;19(10):1877-1884. [doi: 10.1089/jwh.2009.1783] [Medline: 20831437]
12. Juon HS, Seung-Lee C, Klassen AC. Predictors of regular Pap smears among Korean-American women. Prev Med 2003 Dec;37(6 Pt 1):585-592. [Medline: 14636792]
13. Lee MC. Knowledge, barriers, and motivators related to cervical cancer screening among Korean-American women. A focus group approach. Cancer Nurs 2000 Jun;23(3):168-175. [Medline: 10851767]
14. Juon HS, Choi Y, Kim MT. Cancer screening behaviors among Korean-American women. Cancer Detect Prev 2000;24(6):589-601. [Medline: 11198273]
15. Ponce NA, Chawla N, Babey SH, Gatchell MS, Etzioni DA, Spencer BA, et al. Is there a language divide in pap test use? Med Care 2006 Nov;44(11):998-1004. [doi: 10.1097/01.mlr.0000233676.61237.ef] [Medline: 17063131]
16. Lee EE, Eun Y, Lee SY, Nandy K. Age-related differences in health beliefs regarding cervical cancer screening among Korean American women. J Transcult Nurs 2012 Jul;23(3):237-245 [FREE Full text] [doi: 10.1177/1043659612441015] [Medline: 22477716]
17. Im EO, Park YS, Lee EO, Yun SN. Korean women's attitudes toward breast cancer screening tests. Int J Nurs Stud 2004 Aug;41(6):583-589. [doi: 10.1016/j.ijnurstu.2004.01.004] [Medline: 15240081]
18. Fang CY, Ma GX, Tan Y, Chi N. A multifaceted intervention to increase cervical cancer screening among underserved Korean women. Cancer Epidemiol Biomarkers Prev 2007 Jun;16(6):1298-1302 [FREE Full text] [doi: 10.1158/1055-9965.EPI-07-0091] [Medline: 17548702]
19. Moskowitz JM, Kazinets G, Wong JM, Tager IB. "Health is strength": a community health education program to improve breast and cervical cancer screening among Korean American Women in Alameda County, California. Cancer Detect Prev 2007;31(2):173-183. [doi: 10.1016/j.cdp.2007.02.002] [Medline: 17418978]
20. Maxwell AE, Jo AM, Chin SY, Lee KS, Bastani R. Impact of a print intervention to increase annual mammography screening among Korean American women enrolled in the National Breast and Cervical Cancer Early Detection Program. Cancer Detect Prev 2008;32(3):229-235. [doi: 10.1016/j.cdp.2008.04.003] [Medline: 18799271]
21. McDonnell DD, Lee HJ, Kim YB, Kazinets G, Moskowitz JM. Cancer coverage in a mainstream and Korean American online newspaper: lessons for community intervention. Patient Educ Couns 2008 Jun;71(3):388-395. [doi: 10.1016/j.pec.2008.03.004] [Medline: 18406097]
22. U.S. Department of Health and Human Services. Health information technology and quality improvement: what is mHealth? URL: http://www.hrsa.gov/healthit/mhealth.html [accessed 2014-08-22] [WebCite Cache ID 6S1KhvpHK]

23. Patrick K, Raab F, Adams M, Dillon L. mDiet: A personalized approach to weight management using text messaging. In: Fogg BJ, Adler R, editors. Texting 4 Health: A Simple, Powerful Way to Change Lives. Palo Alto, CA: Stanford University; 2009:35-47.
24. Smith M, Harris L. Text "Ahhhh" for me please. In: Fogg BJ, Adler R, editors. Texting 4 Health. Palo Alto, CA: Stanford University; 2009:49-56.
25. Riley W, Obermayer J, Jean-Mary J. Internet and mobile phone text messaging intervention for college smokers. J Am Coll Health 2008;57(2):245-248. [doi: 10.3200/JACH.57.2.245-248] [Medline: 18809542]
26. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17 Suppl 1:44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]
27. 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]
28. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based behavior change or disease management interventions for health care consumers: A systematic review. PLos Med 2013;10(1):e1001362. [doi: 10.1371/journal.pmed.1001362]
29. Levine D. Using technology to promote youth sexual health. In: Fogg BJ, Adler R, editors. Texting 4 Health. Palo Alto, CA: Stanford University; 2009:57-65.
30. Fattor J, Friedlander A. Texting 2 increase physical activity. In: Fogg BJ, Adler R, editors. Texting 4 Health. Palo Alto, CA: Stanford University; 2009:67-74.
31. Mehta R. Zume life: Creating self-care solutions. In: Fogg BJ, Adler R, editors. Texting 4 Health. Palo Alto, CA: Stanford University; 2009:75-84.
32. Ryan D, Cobern W, Wheeler J, Price D, Tarassenko L. Mobile phone technology in the management of asthma. J Telemed Telecare 2005;11 Suppl 1:43-46. [doi: 10.1258/1357633054461714] [Medline: 16035991]
33. Fogg BJ. A behavior model for persuasive design. In: Commun ACM. 2009 Presented at: 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, California URL: http://bjfogg.com/fbm_files/page4_1.pdf
34. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
35. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
36. Champion VL. Instrument refinement for breast cancer screening behaviors. Nurs Res 1993;42(3):139-143. [Medline: 8506161]
37. Taylor VM, Jackson JC, Tu SP, Yasui Y, Schwartz SM, Kuniyuki A, et al. Cervical cancer screening among Chinese Americans. Cancer Detect Prev 2002;26(2):139-145 [FREE Full text] [Medline: 12102148]
38. Khokhar A. Short text messages (SMS) as a reminder system for making working women from Delhi Breast Aware. Asian Pac J Cancer Prev 2009;10(2):319-322 [FREE Full text] [Medline: 19537904]
39. PEW Research Center. Mobile Technology Fact Sheet URL: http://www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ [accessed 2014-07-18] [WebCite Cache ID 6R9O9QWcn]
40. Miller BA, Chu KC, Hankey BF, Ries LA. Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the U.S. Cancer Causes Control 2008 Apr;19(3):227-256 [FREE Full text] [doi: 10.1007/s10552-007-9088-3] [Medline: 18066673]
41. Wang SS, Carreon JD, Gomez SL, Devesa SS. Cervical cancer incidence among 6 Asian ethnic groups in the United States, 1996 through 2004. Cancer 2010 Feb 15;116(4):949-956 [FREE Full text] [doi: 10.1002/cncr.24843] [Medline: 20029972]
42. Mills PK, Yang RC, Riordan D. Cancer incidence in the Hmong in California, 1988-2000. Cancer 2005 Dec 15;104(12 Suppl):2969-2974 [FREE Full text] [doi: 10.1002/cncr.21525] [Medline: 16247793]

Abbreviations

FBM: Fogg’s Behavior Model
MMS: multimedia message service
mScreening: mobile phone text-message-based Screening
SMS: short message service

Edited by G Eysenbach; submitted 02.06.14; peer-reviewed by K Muessig, C Harrington; comments to author 16.07.14; accepted 28.07.14; published 27.08.14.

Please cite as:
Lee HY, Koopmeiners JS, Rhee TG, Raveis VH, Ahluwalia JS
Mobile Phone Text Messaging Intervention for Cervical Cancer Screening: Changes in Knowledge and Behavior Pre-Post Intervention
J Med Internet Res 2014;16(8):e196
URL: http://www.jmir.org/2014/8/e196/
doi:10.2196/jmir.3576
PMID:25164545

Original Paper

Representation of Health Conditions on Facebook: Content Analysis and Evaluation of User Engagement

Timothy M Hale[1], PhD; Akhilesh S Pathipati[2], BA; Shiyi Zan[1], BS; Kamal Jethwani[1], MD, MPH

[1]Center for Connected Health, Partners HealthCare, Inc, Harvard Medical School, Boston, MA, United States
[2]Stanford University School of Medicine, Stanford, CA, United States

Corresponding Author:
Timothy M Hale, PhD
Center for Connected Health
Partners HealthCare, Inc
Harvard Medical School
25 New Chardon Street, Suite 300
Boston, MA, 02114
United States
Phone: 1 617 643 9852
Fax: 1 617 726 7530
Email: tmhale@mgh.harvard.edu

Abstract

Background: A sizable majority of adult Internet users report looking for health information online. Social networking sites (SNS) like Facebook represent a common place to seek information, but very little is known about the representation and use of health content on SNS.

Objective: Our goal in this study was to understand the role of SNS in health information seeking. More specifically, we aimed to describe how health conditions are represented on Facebook Pages and how users interact with these different conditions.

Methods: We used Google Insights to identify the 20 most searched for health conditions on Google and then searched each of the resulting terms on Facebook. We compiled a list of the first 50 Facebook "Pages" results for each health condition. After filtering results to identify pages relevant to our research, we categorized pages into one of seven categories based on the page's primary purpose. We then measured user engagement by evaluating the number of "Likes" for different conditions and types of pages.

Results: The search returned 50 pages for 18 of the health conditions, but only 48 pages were found for "anemia" and 5 pages were found for "flu symptoms", yielding a total of 953 pages. A large number of pages (29.4%, 280/953) were irrelevant to the health condition searched. Of the 673 relevant pages, 151 were not in English or originated outside the United States, leaving 522 pages to be coded for content. The most common type of page was marketing/promotion (32.2%, 168/522) followed by information/awareness (20.7%, 108/522), Wikipedia-type pages (15.5%, 81/522), patient support (9.4%, 49/522), and general support (3.6%, 19/522). Health conditions varied greatly by the primary page type. All health conditions had some marketing/promotion pages and this made up 76% (29/38) of pages on acquired immunodeficiency syndrome (AIDS). The largest percentage of general support pages were cancer (19%, 6/32) and stomach (16%, 4/25). For patient support, stroke (67%, 4/6), lupus (33%, 10/30), breast cancer (19%, 6/31), arthritis (16%, 6/36), and diabetes (16%, 6/37) ranked the highest. Six health conditions were not represented by any type of support pages (ie, human papillomavirus, diarrhea, flu symptoms, pneumonia, spine, human immunodeficiency virus). Marketing/promotion pages accounted for 46.73% (10,371,169/22,191,633) of all Likes, followed by support pages (40.66%, 9,023,234/22,191,633). Cancer and breast cancer accounted for 86.90% (19,284,066/22,191,633) of all page Likes.

Conclusions: This research represents the first attempts to comprehensively describe publicly available health content and user engagement with health conditions on Facebook pages. Public health interventions using Facebook will need to be designed to ensure relevant information is easy to find and with an understanding that stigma associated with some health conditions may limit the users' engagement with Facebook pages. This line of research merits further investigation as Facebook and other SNS continue to evolve over the coming years.

(J Med Internet Res 2014;16(8):e182) doi:10.2196/jmir.3275

KEYWORDS

Internet; Facebook; social media; social networking sites; social support; health communication; information seeking behavior

Introduction

The Internet has radically changed how most people find and share information about health and medical conditions. The practice of looking for health information online has become increasingly popular, with 59% of US adults (72% of adult Internet users) reporting that they have done so in the past year [1]. Nearly half of these individuals reported that the information they found online led them to believe they needed to seek health attention [1]. Even for serious health conditions such as cancer, people are more likely to turn to the Internet first for health information despite a greater trust in their doctor as a source of information [2]. There are many reasons for the rise of the Internet as a source of health information, including: 24/7 availability, the greater anonymity it offers for those with sensitive health care needs, and the opportunity to locate and connect with other people with similar health conditions [3].

Social media is a relatively new health communication channel that enables people to communicate and interact with a larger number of people, find and share information about their health and medical conditions, and receive health messages [4]. Social networking sites (SNS) are one of the most popular and widely used forms of social media with 72% of online US adults using SNS, as of May 2013 [5]. Facebook is the most widely used SNS [6], with 93% of online US adult users reporting having a Facebook account [5] and with 727 million daily users worldwide [7]. Facebook began as an online social network for college students and remains popular with young adults—86% of Internet users aged 18-29 years use Facebook [6]. There has been a significant upward trend in its adoption by older adults in recent years and now 73% of Internet users aged 30 to 49 years and 57% of those aged 50 to 64 years report using Facebook [6]. This is especially significant considering that individuals become more likely to develop chronic health conditions as they age.

Despite the rapid and widespread adoption of Facebook among Internet users, little is known about the broader representation of health conditions on Facebook. The existing literature has largely focused on a small number of specific health conditions that have taken one of two approaches, either (1) an aggregated content analysis of posts, or (2) a more detailed analysis of differences in the primary purpose of groups, the number of members, and the content of posts.

Studies examining the content of posts find marked differences by health condition. Greene et al [8] examined diabetes groups and found that two-thirds of wall posts and discussion topics were characterized by sharing of information on diabetes management strategies, followed by posts related to emotional support and promotional themes. In contrast to the findings for diabetes groups, Ahmed et al [9] found that among Facebook groups devoted to concussions, nearly two-thirds of posts were to relate personal experiences of a concussion and posts were only rarely used to seek information (8%) or offer advice (2%). Gajaria et al [10] examined posts by youth to attention deficit hyperactivity disorder (ADHD) Facebook groups. They found the largest percentage of posts (42%) were about defining ADHD and creating a sense of group identity, and to seek and share advice regarding medications and symptom management (35%).

Bender et al and Thoren et al examined types of groups, number of members, and the content of posts. Bender et al [11] examined the content of breast cancer groups and found that most groups were created for fundraising or awareness purposes, rather than supportive care. They also found that the awareness groups had the most members, while the support groups generated the greatest number of posts. Thoren et al [12] examined the content of Facebook groups focusing on premature infants. Similar to findings from breast cancer groups, they found that most premature infant groups were devoted to fundraising or awareness purposes and that these groups had the most members. However, despite the emphasis on fundraising and awareness groups, 53% of all posts were for “interpersonal support” and 31% for “information sharing”.

To our knowledge, only two studies have been conducted to characterize the representation of a broader range of health conditions on Facebook [13,14]. The most comprehensive study was conducted by Farmer et al [14], who constructed a list of search terms based on the 11 most prevalent non-communicable diseases identified by the World Health Organization. Using both health and lay terms, they searched all Facebook groups between December 2007 and January 2009. They found that respiratory groups, diabetes, cardiovascular disease, and digestive disease made up the largest number of groups, while groups related to malignant neoplasms had the most members. Patient groups comprised of disease sufferers were the most common (47%), followed by support (28%), and fundraising (19%) groups. De la Torre-Díez et al [13] examined how three diseases with the greatest public burden (ie, breast cancer, colorectal cancer, and diabetes) are represented on Facebook and Twitter. Conducting a search in 2011, they found that “prevention” groups that seek to raise awareness and/or money of a disease was the most popular categorization for all three diseases (18%), followed by support groups (17.9%), and research investigations (14.3%).

Taken together, these studies have begun to demonstrate how people use Facebook to find and share health information. However, these findings fail to reflect the representation of health conditions on Facebook, due to the focus on specific health conditions or the limited inclusion criteria for disease groups (ie, non-communicable diseases, diseases with greatest public burden). Therefore, we still lack a comprehensive review of how health conditions are represented on Facebook.

In this paper, we aimed to (1) describe the results of a search for 20 common health conditions on Facebook “Pages”, (2) identify the purpose and content of these pages, and (3) evaluate user engagement with these pages.

Results may offer important insights for future public health initiatives. For example, a better understanding of which conditions are prominent on Facebook provides perspective on the accessibility of information on different diseases. Second, variation in accessibility may have further implications for class-specific engagement with health conditions on Facebook. Finally, data on user engagement may provide a means for health professionals to more effectively disseminate information on Facebook.

Methods

Facebook Pages

Unlike previous studies, we chose to focus our search on Facebook “Pages” rather than “Groups”. In the evolution of Facebook, groups initially served as a primary forum for communication and, as such, many of the previous studies cited above conducted their research within this realm. However, the functionality of groups began to shift following the introduction of Facebook Pages in 2007, initially created as a way to allow public profile owners (individuals, organizations, services, etc) to advertise to Facebook users more easily. These “Fan Pages” behaved much like a user’s profile and allowed owners to send updates to those who subscribed to their page and access insights and analytics of their fan base. Until April 19, 2010, users had the option to become a “Fan” of a page; this subsequently changed so that users could “Like” a page. These “Like Pages” allow for an unlimited number of “Likers” and have additional functionalities including the ability to add tabs for email collection and specialized content. In addition, “Community Pages” were also introduced around this time, allowing for the integration of content into Facebook directly from Wikipedia pages.

Pages and groups differ by their function: pages can be thought to resemble a promotional blog, whereas groups are more analogous to a moderated message board. The key benefits of a page over a group is that pages (1) are able to get internal promotion through the page feed of fans after they like a page, (2) have more options for customization, (3) have greater search engine visibility, (4) allow the creator or administrator of the page to remain anonymous, and (5) give the user more power to control the content they receive from the page.

In more recent years, pages have exploded in popularity as a means for publicly accessible interaction, to the point that some social media commentators have even described groups as “obsolete” or as a “Facebook fossil” [15]. Along with this shift, more and more groups have become “closed” (visible on Facebook but content is visible to members only) or “secret” (completely invisible to all on Facebook except for invited members), making information contained within groups no longer easily accessible to a casually browsing Facebook user. Therefore, we decided to focus our study on pages, as we believe that it would be able to provide a more complete picture of a health condition’s representation on Facebook.

Search Criteria and Strategy

On July 24, 2012, we identified the 20 most searched for health conditions on Google using Google Insights (see Table 1). These 20 conditions provided the basis for our subsequent searches on Facebook. On the same day, we conducted searches for these 20 health conditions on Facebook using Facebook Search. For our searches, we specifically focused on Facebook pages, excluding search results for people, groups, and other categories. We recorded the top 50 pages results for each health condition, as well as the URL and the number of Likes each page had received from Facebook members.

The Facebook search algorithm is user-centric and search results will vary for different people based on their past Facebook use, profile information, and network of friends [16]. To minimize this effect, we created a new Facebook account using minimal biographical information, specifying only name, gender, and age: Jonathan Davis, male, 45 years old. By creating a generic profile, we hoped to retrieve search results that would be more representative of health conditions on Facebook and that are not tailored to individual factors, social context, or geographical location of the person searching. We ensured that our new profile had no friends and no preexisting Likes. We also deleted and disabled cookies and location services prior to conducting our searches.

Once we had compiled lists of 50 pages for each search term, we filtered our search results to limit our analysis to those pages that were relevant to the health condition (see Figure 1). For instance, when we conducted our search for human immunodeficiency virus (HIV), many of the top 50 pages were fan pages for a band named “The Hive”. Similarly, many results for diarrhea were for a band named “Raging Diarrhea”, or other topics irrelevant to the health condition. We then further restricted our analysis to pages that were in English, and that were based in the United States. If a country was not specified but the page was in English, we assumed that it was based in the United States. After filtering the pages with our criteria, we generated a list of “clean” pages for each condition, which listed the name of each page and the number of Likes it had received.

Table 1. Google Insights results: top 20 health condition search terms used in the United States between September 2007-June 2012.[a]

Google Insights	Health condition
1	cancer
2	diabetes
3	stomach
4	herpes
5	back pain
6	human immunodeficiency virus (HIV)
7	blood pressure
8	thyroid
9	breast cancer
10	arthritis
11	acquired immunodeficiency syndrome (AIDS)
12	lupus
13	diarrhea
14	pneumonia
15	spine
16	flu symptoms
17	human papilloma virus (HPV)
18	asthma
19	anemia
20	stroke

[a] Search conducted on July 24, 2012.

Figure 1. Search result workflow.

Coding Page Content and Descriptive Analysis

To determine categories for classifying pages, we started with a literature review to identify previous categorizations of groups on Facebook. As noted in the introduction, Facebook pages and groups serve somewhat different roles, but previous research provided a starting point for identifying their purpose. For example, Greene et al [8] used five categories: advertisements, providing information, requesting information, support, and irrelevant; De la Torre-Díez et al [13] identified five categories: fund collecting, awareness, support, prevention, and disease-fighting; Bender et al [11] used four categories: fundraising, awareness, promote-a-site, and support; and Farmer et al [14] used four categories: patient groups, support groups, fundraising/charity groups, and other. Based on this review, we initially chose to cluster pages into five categories: patient support, general support, information/awareness, marketing/promotion, and other.

Once we compiled our list of pages, two co-authors (ASP and SZ) evaluated the 20 most recent posts on each page and categorized page content into one of five types. An example Facebook page is presented in Figure 2. The two coders conducted an initial categorization of approximately 90 pages in order to determine interrater reliability (IRR). Although the IRR was acceptable (Cohen’s kappa=.74), there was disagreement on how to code pages that lacked content, or were Wikipedia-type informational pages with no user content. As a result, we added categories for Wikipedia and blank pages, giving us seven categories (see Table 2). Another 90 pages were coded using the new classification scheme and the IRR improved (Cohen’s kappa=.83). The remaining pages were then divided between the two coders.

Once all pages were coded, data was aggregated for each condition. We first aggregated data on the number of pages by health condition and type of content. We then compiled data on the number of page Likes by condition and calculated the total number of Likes by page content. We used the number of Likes as a proxy for member interest or engagement with a health condition on Facebook.

Figure 2. Screenshot of a Facebook page analyzed in this study.

Breast Cancer Awareness
3,383,182 likes · 82,158 talking about this

Health/Wellness
Click for FREE to fund mammograms @ www.TheBreastCancerSite.com

Breast Cancer Awareness shared a link.
9 hours ago

Click for Dr. Harness' tips on how to face your diagnosis without becoming overwhelmed.

I Have Breast Cancer: Is it Fatal?
answers.thebreastcancersite.com
The word "cancer" strikes fear in every heart, but a diagnosis can mean a variety of things.

211 people like this.

Breast Cancer Awareness
12 hours ago

~❤~ Pink Ribbon Earrings in 10K Pink Gold – Crafted from solid Victorian-style 10K pink gold! ~❤~

Likes

- **Pink Ribbon App** App Page
- **The Breast Cancer Site Store** 2 friends also like this.
- **The Breast Cancer Site** 4 friends also like this.
- **The Pink Ribbon** App Page
- **Support The Fight Against Breast Cancer** Non-Profit Organization

Breast Cancer Awareness shared a link.
23 hours ago

"More than a year after my last treatment, I still have the amazing support of my friends and the same positive attitude I had on that first tough day..."

---▶ LIKE if you believe that support and a positive attitude truly can make all the difference!

READ: Friends, Support, & Laughter
supportthefightagainstbreastcancer.com
The day I was diagnosed, my family and friends were there to support me...instantly! When I told one of my best friends, De, the first thing she said was "I'll shave my head

Sponsored

DIA 2012 Meeting
Join over 7500 professionals involved in the discovery and development of pharmaceuticals.
1,598 people like DIA Europe.

Rent Your Roof For Cash!
1bog.org
NEWS BRIEF: Companies are paying homeowners to go solar. Sign up now & find out how!

Table 2. Page categorization and example posts.

Categorization	Description	Example page and post
1. Patient Support	Characterized by emotional and informational support for patients of the condition. Often included motivational messages, links, and posts by affected individuals.	Kick Cancer “No matter how scary things may seem or how bad they may look, keep going and never give up! #kickcancer”
2. General Support	Characterized by emotional and informational support for caregivers, family, friends, and some patients themselves. Often included motivational messages and posts by supporters of affected individuals.	Mesothelioma Cancer Alliance “I hope they find a way to diagnose this disease sooner so more people can be cured and not have to go thru the devastation so many of us have had to go thru. Rest in peace daddy.”
3. Marketing/Promotion	Characterized by promotion of specific products, events, or institutions. Included self-promotion of the page or events and organization run by the page managers.	Cancer Sucks “New Cancer Sucks Purple Heart design tee’s available in both men’s and women’s styles! Get yours now!”
4. Information/Awareness	Characterized by a focus on raising awareness of a condition or facilitating information exchange. Included many links to information, treatment recommendations, and research pages.	Breast Cancer Awareness “medicalxpress.com: Accelerated radiation treatment effective for noninvasive breast cancer”
5. Wikipedia	Some pages simply provided information from Wikipedia on the condition.	Cancer “From Wikipedia, the free encyclopedia”
6. Blank	Blank pages were pages that addressed the condition, but did not have any posts.	Diarrhea None
7. Other	Any pages that did not fit the above categories were classified as “other.” For example, we classified personal blogs by people who were affected by a condition as “other”.	Mrs Lupus “Here is one of my most popular blog posts, it is not very informative but it is written with raw emotion that all us lupies can understand.”

Results

Facebook Page Search

Our first aim was to describe the results of a search for 20 common health conditions on Facebook. We used the Facebook search function to list the first 50 pages found for each of the 20 health conditions identified using Google Insights (see Table 3). The search returned 50 pages for 18 of the health conditions, but only 48 pages were found for “anemia” and five pages were found for “flu symptoms”. Thus, the list or sample of pages returned was 953 pages.

The Facebook search turned up a number of irrelevant pages that were not about health conditions. Of the 953 pages returned in the search, 280 pages were not about the health condition used in the search term. The number of relevant pages also varied considerably by health condition. Conditions with the greatest number of relevant pages were breast cancer and diabetes (n=50), followed by cancer, thyroid, and arthritis (n=49). Conditions with the lowest number of relevant pages were stroke (n=10), HIV (n=10), spine (n=18), human papillomavirus (HPV; n=23), and diarrhea (n=23). The search for flu symptoms yielded only five pages, but 100% of the pages were relevant.

A second criteria was that pages be in English and have a user base located in the United States. This further reduced the number of relevant pages by 151 to 522 pages. Five conditions (stroke, HPV, asthma, breast cancer, and cancer) had one-third or more of relevant pages in a language other than English or located outside North America. The median number of relevant pages for each health condition was 29 and ranged from 5 to 43 pages (results not shown in Table 3).

Table 3. Relevant Facebook pages.

Google Insights ranking	Health condition	Pages sampled	Relevant, n (%)	Eliminated non-US, n (%)[a]	Clean pages[b], n (%)[c]
1	cancer	50	49 (98.0)	17 (34.7)	32 (64.0)
2	diabetes	50	50 (100.0)	13 (26.0)	37 (74.0)
3	stomach	50	26 (52.0)	1 (3.9)	25 (50.0)
4	herpes	50	33 (66.0)	5 (15.2)	28 (56.0)
5	back pain	50	44 (88.0)	10 (22.7)	34 (68.0)
6	human immunodeficiency virus (HIV)	50	10 (20.0)	2 (20.0)	8 (16.0)
7	blood pressure	50	43 (86.0)	6 (14.0)	37 (74.0)
8	thyroid	50	49 (98.0)	6 (12.2)	43 (86.0)
9	breast cancer	50	50 (100.0)	19 (38.0)	31 (62.0)
10	arthritis	50	49 (98.0)	13 (26.5)	36 (72.0)
11	acquired immunodeficiency syndrome (AIDS)	50	43 (86.0)	5 (11.6)	38 (76.0)
12	lupus	50	37 (74.0)	7 (18.9)	30 (60.0)
13	diarrhea	50	23 (46.0)	4 (17.4)	19 (38.0)
14	pneumonia	50	39 (78.0)	3 (7.7)	36 (72.0)
15	spine	50	18 (36.0)	4 (22.2)	14 (28.0)
16	flu symptoms	5	5 (100.0)	0 (0.0)	5 (100.0)
17	human papillomavirus (HPV)	50	23 (46.0)	9 (39.1)	14 (28.0)
18	asthma	50	41 (82.0)	16 (39.0)	25 (50.0)
19	anemia	48	31 (64.6)	7 (22.6)	24 (50.0)
20	stroke	50	10 (20.0)	4 (40.0)	6 (12.0)
Totals		953	673 (70.6)	151 (22.4)	522 (54.8)

[a] Percent of health condition relevant pages.

[b] Mean number of pages=26; median number of pages=29; interquartile range=18.25.

[c] Percent of pages sampled.

Health Conditions on Facebook Pages

Our second aim was to identify the content of the 522 pages identified as relevant to health conditions. The most frequent page type was marketing/promotion, which accounted for 168/522 or 32.2% of pages. The next most frequent page types were information and awareness of a health condition (20.7%, 108/522) followed by Wikipedia-type pages (15.5%, 81/522), patient support (9.4%, 49/522), and general support (3.6%, 19/522). A total of 64 pages were coded as “other” type. Finally, 33 pages that did not contain enough information were coded as blank.

Next, we examined pages by health condition and content (see Tables 4 and 5). All health conditions had some pages devoted to marketing/promotion, ranging from 76% (29/38) of acquired immunodeficiency syndrome (AIDS) pages to 5% (1/19) of diarrhea pages. Six conditions (ie, AIDS, arthritis, spine, breast cancer, asthma, and cancer) had more than 40% of the pages devoted primarily to marketing/promotion. Wikipedia pages comprised a large percentage of acute conditions (anemia, 58%, 14/24; pneumonia, 58%, 21/36; flu symptoms, 40%, 2/5; and diarrhea, 26%, 5/19) but formed a much smaller proportion for breast cancer (3%, 1/31) and cancer (3%, 1/32) pages. For HIV, spine, and diabetes pages, 40% or more were information/awareness-type pages. In contrast to the high percentage of Wikipedia pages devoted to pneumonia, diarrhea, and flu symptoms, these conditions made up a small percentage of information/awareness pages. Cancer and breast cancer made up 25% (8/32) and 23% (7/31) of information/awareness pages, respectively. Conditions with the largest percentage of support pages were cancer (19%, 6/32) and stomach (16%, 4/25) for general support, and stroke (67%, 4/6), lupus (33%, 10/30), breast cancer (19%, 6/31), arthritis (16%, 6/36), and diabetes (16%, 6/37) for patient support. A large number of health conditions were not represented by any type of support pages (ie, HPV, diarrhea, flu symptoms, pneumonia, spine, HIV). Over 30% of blood pressure, diarrhea, and herpes pages could not be classified as one of the other types and were categorized under other. Over 25% of HPV and diarrhea pages did not contain sufficient information to determine the purpose of the pages and were categorized as blank.

Table 4. Facebook pages by health condition and content—General Support, Patient Support, and Information/Awareness.

Google Insights ranking	Health condition	Clean pages, n	General support, n (%)	Patient support, n (%)	Information, n (%)
1	cancer	32	6 (18.8)	3 (9.4)	8 (25.0)
2	diabetes	37	1 (2.7)	6 (16.2)	15 (40.5)
3	stomach	25	4 (16.0)	1 (4.0)	7 (28.0)
4	herpes	28	1 (3.6)	0 (0.0)	4 (14.3)
5	back pain	34	0 (0.0)	3 (8.8)	11 (32.4)
6	human immunodeficiency virus (HIV)	8	0 (0.0)	0 (0.0)	4 (50.0)
7	blood pressure	37	0 (0.0)	1 (2.7)	10 (27.0)
8	thyroid	43	0 (0.0)	5 (11.6)	11 (25.6)
9	breast cancer	31	1 (3.2)	6 (19.4)	7 (22.6)
10	arthritis	36	1 (2.8)	6 (16.7)	3 (8.3)
11	acquired immunodeficiency syndrome (AIDS)	38	3 (7.9)	0 (0.0)	3 (7.9)
12	lupus	30	1 (3.3)	10 (33.3)	5 (16.7)
13	diarrhea	19	0 (0.0)	0 (0.0)	1 (5.3)
14	pneumonia	36	0 (0.0)	0 (0.0)	2 (5.6)
15	spine	14	0 (0.0)	0 (0.0)	6 (42.9)
16	flu symptoms	5	0 (0.0)	0 (0.0)	0 (0.0)
17	human papillomavirus (HPV)	14	0 (0.0)	0 (0.0)	2 (14.3)
18	asthma	25	1 (4.0)	2 (8.0)	7 (28.0)
19	anemia	24	0 (0.0)	2 (8.3)	2 (8.3)
20	stroke	6	0 (0.0)	4 (66.7)	0 (0.0)
Total		522	19 (3.6)	49 (9.4)	108 (20.7)

Table 5. Facebook pages by health condition and content—Wikipedia, Marketing, Other, and Blank.

Google Insights ranking	Health condition	Clean pages, n	Wikipedia, n (%)	Marketing, n (%)	Other, n (%)	Blank, n (%)
1	cancer	32	1 (3.1)	13 (40.6)	0 (0.0)	1 (3.1)
2	diabetes	37	3 (8.1)	11 (29.7)	1 (2.7)	0 (0.0)
3	stomach	25	2 (8.0)	3 (12.0)	5 (20.0)	3 (12.0)
4	herpes	28	5 (17.9)	5 (17.9)	9 (32.1)	4 (14.3)
5	back pain	34	2 (5.9)	10 (29.4)	6 (17.7)	2 (5.9)
6	human immunodeficiency virus (HIV)	8	2 (25.0)	2 (25.0)	0 (0.0)	0 (0.0)
7	blood pressure	37	5 (13.5)	4 (10.8)	14 (37.8)	3 (8.1)
8	thyroid	43	7 (16.3)	9 (20.9)	8 (18.6)	3 (7.0)
9	breast cancer	31	1 (3.2)	16 (51.6)	0 (0.0)	0 (0.0)
10	arthritis	36	3 (8.3)	22 (61.1)	1 (2.8)	0 (0.0)
11	acquired immunodeficiency syndrome (AIDS)	38	2 (5.3)	29 (76.3)	1 (2.6)	0 (0.0)
12	lupus	30	0 (0.0)	11 (36.7)	3 (10.0)	0 (0.0)
13	diarrhea	19	5 (26.3)	1 (5.3)	7 (36.8)	5 (26.3)
14	pneumonia	36	21 (58.3)	2 (5.6)	4 (11.1)	7 (19.4)
15	spine	14	0 (0.0)	8 (57.1)	0 (0.0)	0 (0.0)
16	flu symptoms	5	2 (40.0)	1 (20.0)	1 (20.0)	1 (20.0)
17	human papillomavirus (HPV)	14	4 (28.6)	3 (21.4)	1 (7.1)	4 (28.6)
18	asthma	25	1 (4.0)	12 (48.0)	2 (8.0)	0 (0.0)
19	anemia	24	14 (58.3)	5 (20.9)	1 (4.2)	0 (0.0)
20	stroke	6	1 (16.7)	1 (16.7)	0 (0.0)	0 (0.0)
Total		522	81 (15.5)	168 (32.2)	64 (12.3)	33 (6.3)

Likes on Facebook Pages

Our third aim was to examine the level of user engagement for each of the 20 health conditions represented on Facebook pages. For each of the conditions, we aggregated the number of Likes for the pages of each condition, as well as the average, median, minimum, and maximum (see Table 6). For the 20 health conditions we searched for on Facebook pages, there were 22,191,633 Likes. The mean number of Likes across all health condition pages was 1,110,240 and ranged from 0 to 3,537,360. Cancer and breast cancer together account for 86.90% (19,284,066/22,191,633) of total Likes. AIDS and diabetes each account for about 4.5% of total Likes, followed by HIV and lupus with about 1.1% each. The remaining 14 health conditions represent less than 2% of the total Likes.

Likes were most often given to marketing/promotion pages, which accounted for 46.73% (10,371,169/22,191,633) of all Likes. Support pages accounted for 40.66% (9,023,234/22,191,633) of total Likes with general support accounting for 35.89% (7,964,328/22,191,633) and patient support for 4.77% (1,058,906/22,191,633). This is in contrast to the findings for the number of pages, in which patient support and general support accounted for relatively small percentages of the total pages (9.4%, 49/522 and 3.6%, 19/522, respectively) compared to information and Wikipedia pages (20.7%, 108/522 and 15.5%, 81/522, respectively).

Finally, we examined how the number of Likes by health condition and type of page content (see Tables 7 and 8). Twelve health conditions were primarily represented by Likes on marketing/promotion pages, with the percentage of pages coded as marketing/promotion exceeding the percentage for any other type (herpes, HIV, HPV, AIDS, flu symptoms, anemia, cancer, lupus, breast cancer, spine, blood pressure, and back pain). A total of 80% or more of herpes, HIV, and HPV Likes were on pages coded as marketing/promotion. Seven conditions were primarily represented by Likes on information/awareness or Wikipedia pages. Pneumonia, diabetes, arthritis, thyroid, and stomach were primarily represented by Likes on information/awareness pages. Only 9-10% of cancer and breast cancer Likes were on information/awareness pages. Cancer and breast cancer Likes were nearly evenly divided between marketing/promotion and combined support pages. Cancer and breast cancer pages accounted for most of the general support Likes (94.59%, 7,533,563/7,964,328) and, although a modest percentage of breast cancer and cancer pages were categorized as patient support (19%, 6/31 and 9%, 3/32 respectively), the large number of Likes for these two health conditions comprised 82.78% (876,589/1,058,906) of the total number of patient support Likes. Wikipedia pages made up the largest percentage of diarrhea and asthma Likes. The largest percentage of Likes on thyroid pages were on those classified as patient support.

Table 6. Facebook “Likes” by health condition.

Google Insights ranking	Health condition	Total Likes, n (%)	Average	Minimum	Maximum	Median
1	cancer	10,228,611 (46.09)	319,644	956	3,537,341	29,776
2	diabetes	986,868 (4.45)	26,672	618	473,585	4529
3	stomach	28,620 (0.13)	1145	202	6044	572
4	herpes	46,778 (0.21)	1671	3	36,730	104
5	back pain	15,165 (0.07)	446	1	2299	282
6	human immunodeficiency virus (HIV)	249,468 (1.12)	31,184	246	207,098	4215
7	blood pressure	9711 (0.04)	262	0	1629	27
8	thyroid	75,539 (0.34)	1757	2	14,249	815
9	breast cancer	9,055,455 (40.81)	292,111	1,005	3,537,360	19,054
10	arthritis	106,565 (0.48)	2960	200	34,578	708
11	acquired immunodeficiency syndrome (AIDS)	1,014,419 (4.57)	26,695	651	551,888	3487
12	lupus	241,286 (1.09)	8043	231	76,297	1712
13	diarrhea	6215 (0.03)	327	2	2329	44
14	pneumonia	6952 (0.03)	193	0	4405	3
15	spine	38,213 (0.17)	2569	197	13,196	588
16	flu symptoms	43 (0.00)	9	0	27	2
17	human papillomavirus (HPV)	5371 (0.02)	384	0	3594	57
18	asthma	24,890 (0.11)	996	192	8279	390
19	anemia	11,887 (0.05)	495	7	2845	190
20	stroke	39,577 (0.18)	6596	2129	18,670	3915
Total		22,191,633 (100.00)				

Table 7. Facebook “Likes” by health condition and content—General Support, Patient Support, and Information/Awareness.

Google Insights ranking	Health condition	Total Likes	General support, n (%)	Patient support, n (%)	Information, n (%)
1	cancer	10,228,611	3,996,203 (39.07)	372,880 (3.65)	905,348 (8.85)
2	diabetes	986,868	207,285 (21.00)	71,482 (7.24)	559,877 (56.73)
3	stomach	28,620	6466 (22.59)	2022 (7.06)	8439 (29.49)
4	herpes	46,778	32 (0.07)	0 (0.00)	3772 (8.06)
5	back pain	15,165	0 (0.00)	1567 (10.33)	3662 (24.15)
6	human immunodeficiency virus (HIV)	249,468	0 (0.00)	0 (0.00)	28,815 (11.55)
7	blood pressure	9711	0 (0.00)	992 (10.22)	3868 (39.83)
8	thyroid	75,539	0 (0.00)	4700 (6.22)	22,761 (30.13)
9	breast cancer	9,055,455	3,537,360 (39.06)	503,709 (5.56)	888,712 (9.81)
10	arthritis	106,565	2124 (1.99)	30,204 (28.34)	42,929 (40.28)
11	acquired immunodeficiency syndrome (AIDS)	1,014,419	214,308 (21.13)	0 (0.00)	19,297 (1.90)
12	lupus	241,286	231 (0.10)	37,287 (15.45)	79,861 (33.10)
13	diarrhea	6215	0 (0.00)	0 (0.00)	1433 (23.06)
14	pneumonia	6952	0 (0.00)	0 (0.00)	4455 (64.08)
15	spine	38,213	0 (0.00)	0 (0.00)	22,885 (59.89)
16	flu symptoms	43	0 (0.00)	0 (0.00)	0 (0.00)
17	human papillomavirus (HPV)	5371	0 (0.00)	0 (0.00)	121 (2.25)
18	asthma	24,890	319 (1.28)	1460 (5.87)	7927 (31.85)
19	anemia	11,887	0 (0.00)	101 (0.85)	1240 (10.43)
20	stroke	39,577	0 (0.00)	32,502 (82.12)	0 (0.00)
Total		22,191,633	7,964,328 (35.87)	1,058,906 (4.77)	2,605,402 (11.73)

Table 8. Facebook “Likes” by health condition and content—Wikipedia, Marketing, Other, and Blank.

Google Insights ranking	Health condition	Total Likes	Wikipedia, n (%)	Marketing, n (%)	Other, n (%)	Blank, n (%)
1	cancer	10,228,611	20,428 (0.20)	4,932,796 (48.23)	0 (0.00)	956 (0.01)
2	diabetes	986,868	36,875 (3.74)	110,731 (11.22)	618 (0.06)	0 (0.00)
3	stomach	28,620	2232 (7.80)	2868 (10.02)	5387 (18.82)	1206 (4.21)
4	herpes	46,778	1412 (3.02)	40,495 (86.57)	810 (1.73)	257 (0.55)
5	back pain	15,165	2726 (17.98)	4701 (31.00)	364 (2.40)	2145 (14.14)
6	human immunodeficiency virus (HIV)	249,468	8966 (3.59)	211,687 (84.86)	0 (0.00)	0 (0.00)
7	blood pressure	9711	1320 (13.59)	3023 (31.13)	504 (5.19)	4 (0.04)
8	thyroid	75,539	9313 (12.33)	19,883 (26.32)	18,850 (24.95)	32 (0.04)
9	breast cancer	9,055,455	1757 (0.02)	4,123,917 (45.54)	0 (0.00)	0 (0.00)
10	arthritis	106,565	9021 (8.47)	17,782 (16.69)	4505 (4.23)	0 (0.00)
11	acquired immunodeficiency syndrome (AIDS)	1,014,419	25,327 (2.50)	751,491 (74.08)	3996 (0.39)	0 (0.00)
12	lupus	241,286	0 (0.00)	114,832 (47.59)	9075 (3.76)	0 (0.00)
13	diarrhea	6215	2436 (39.20)	321 (5.16)	1956 (31.47)	69 (1.11)
14	pneumonia	6952	1705 (24.53)	605 (8.70)	167 (2.40)	20 (0.29)
15	spine	38,213	0 (0.00)	15,328 (40.11)	0 (0.00)	0 (0.00)
16	flu symptoms	43	13 (30.23)	27 (62.79)	1 (2.33)	2 (4.65)
17	human papillomavirus (HPV)	5371	790 (14.71)	4329 (80.60)	7 (0.13)	124 (2.31)
18	asthma	24,890	8279 (33.26)	5134 (20.63)	1771 (7.12)	0 (0.00)
19	anemia	11,887	4103 (34.52)	6273 (52.77)	170 (1.43)	0 (0.00)
20	stroke	39,577	2129 (5.38)	4946 (12.50)	0 (0.00)	0 (0.00)
Total		22,191,633	138,832 (0.63)	10,371,169 (46.71)	48,181 (0.22)	4815 (0.02)

Discussion

Principal Findings

We used Google Insights to identify the 20 most searched for health conditions and then searched for these terms on Facebook. We found that a large number of pages were not about the health condition searched, but a similarly named topic. The most common type of page content was marketing/promotion, followed by information/awareness. Only a small number of pages were devoted to social support and six conditions were not represented by any support pages (ie, HPV, diarrhea, flu symptoms, pneumonia, spine, HIV). We also found that engagement measured by Likes was greater for general support and marketing/promotion than for patient support and information/awareness pages.

Relevant Pages

A Facebook search for health conditions returned a large number of page results that were not relevant to the health condition searched (29.4%, 280/953). Additionally, the percentage of relevant pages varies considerably by health condition. While 98% or more of pages listed for six conditions were relevant (flu symptoms, diabetes, breast cancer, cancer, thyroid, arthritis), less than 50% of pages were relevant for five conditions (HIV, stroke, spine, HPV, diarrhea; see Tables 4 and 5).

The variation in the number of relevant pages may be due to the breadth of health conditions that we searched for and/or the method used to identify the Facebook groups and pages. Previous research examining Facebook groups for specific health conditions found that most content was relevant. A total of 97% of the posts were relevant on the 25 largest Facebook groups, focusing on premature infants [12] and on the largest Facebook diabetes groups [8]. Ahmed and colleagues [9] found that 89% of posts on 17 Facebook groups related to concussions were relevant. In contrast, using a search method similar to the one we used, Sajadi and Goldman [17] examined the usefulness of the first 30 listed results for the search term “incontinence” on Facebook, Twitter, and YouTube. They found that nearly half of the search results on Facebook led to pages with no useful information. This problem may be overrepresented in our study due to our search methodology of using a “clean” Facebook profile. With more information about a user, Facebook is likely to show pages that are more relevant to the user, which may also be more relevant to the condition searched.

The difficulty in finding Facebook pages with relevant health information may pose a significant barrier for people with inadequate digital skills. A growing body of literature finds that people with better Internet skills are more likely to go online to search for information, including health information [18,19], and make more varied and effective use of online information resources [20]. Additionally, Internet and information seeking skills vary by socioeconomic status and prior Internet access and use [19,20]. These digital inequalities may limit the utility of Facebook as a health communication channel for people from socially disadvantaged groups and may in fact contribute to increasing knowledge gaps [21-24] and health disparities. Public health interventions that use Facebook as a health communication channel will need to be designed to ensure that information is easy to find for all members of the target population.

Page Content and Social Support

One benefit of using social media for health communication, identified by Moorhead et al [4] in a systematic review of 98 research articles, is the ability for people to draw social support from a large network of friends, relatives, and other users. We found, however, that only 13.0% (68/522) of pages were devoted to social support and that the largest percentage of pages were marketing/promotion (32.2%, 168/522) and information (20.7%, 108/522). Additionally, the percentage of social support pages varied considerably by health condition. For example, several health conditions were represented by few or no social support pages (HIV, AIDS, HPV, herpes, diarrhea, flu symptoms, pneumonia, anemia, blood pressure, and spine) and were largely represented by information and marketing/promotion pages. In contrast, five health conditions (stroke, lupus, cancer, breast cancer, stomach) were represented by 20-67% by social support pages.

Direct comparisons with other studies are difficult due to the differences in the focus on pages versus groups, classification schemes, and the range of health conditions examined. However, the relative lack of pages devoted to social support that we found is consistent with the findings of Bender et al [11] of Facebook breast cancer groups. Although we found a greater percentage of breast cancer pages devoted to social support (22.6%, 7/31) than Bender et al among groups (7%), they found that groups devoted to fundraising (45%) and raising awareness (38%) were most common. In contrast, Farmer et al [14] found that support groups made up a substantial percentage of groups for the 11 most prevalent non-communicable diseases on Facebook. They found that patient groups accounted for 47% of groups, followed by patient/caregiver support groups (28%), and fundraising groups (19%).

The relatively low percentage of social support pages for some health conditions may be due to the higher level of stigma associated with these conditions (ie, HIV, AIDS, HPV, herpes) compared to non-communicable diseases (ie, stroke, lupus, cancer, breast cancer). Rains [25] found that anonymity was one strategy used by people who are embarrassed by their illness and that people with high levels of online anonymity disclosed more health experiences. The lack of anonymity on Facebook may pose a barrier to people's willingness to disclose information about their health condition or to provide open support to other users and limit the effectiveness of public health interventions for some health conditions. Further research is needed to explore how perceived stigma and illness-related embarrassment influences people's willingness to disclose information and express social support.

Engagement

Third, we found that engagement measured by Likes was disproportionate to the number of pages in each category. For example, general support and marketing/promotion pages comprised a larger percentage of Likes than would be expected given the percentage of pages. General support pages represented only 3.6% (19/522) of pages but comprised 35.89% (7,964,328/22,191,633) of Likes. Similarly, marketing/promotion represented 32.2% (168/522) of pages but comprised 46.73% (10,371,169/22,191,633) of Likes. In contrast, patient support and information/awareness pages were underrepresented in Likes compared to percentage of pages, while Wikipedia pages received no Likes.

The greater engagement with general support and marketing/promotion pages versus patient support and information/awareness pages may have to do with what Facebook users view as appropriate use and activities on the site. Lampe et al [26] studied the perception of Facebook's value as an information source and found that on average, users did not find it appropriate to seek information on Facebook and were not likely to make extensive use of the site for information seeking. Another reason may be that marketing/promotion pages have commercial interest in gaining popularity and may employ methods like Facebook ads or viral campaigns to increase the visibility and engagement with their pages through likes. Future research should explore users' perceived norms surrounding the use of Facebook and how these norms impact disclosure of health conditions, seeking health information, and providing social support to people who are ill.

Limitations and Future Research

Although this study has several strengths, including the examination of the search results for 20 health conditions and nearly 1000 Facebook pages and a content categorization scheme based on previous health communication research, it has several limitations.

First, in the time since we collected our data, Facebook has modified its search function. Starting in January 2013, Facebook began to roll out a new "graph search", which became available to all of Facebook's English (US) users by the end of July. The new search function includes three visible changes. First, search results are formatted slightly differently: profile pictures and fonts are larger and more prominent. Second, the search results now include a column on the right side of the page, which features the name, profile picture, and cover picture for the top search result as well as Web searches for the search term. Finally, search listings now include a line for each page that indicates pages that "people also like."

In practice, these changes appear to have little impact on searches regarding health conditions. Even so, the new search may have network effects that will impact future search results and which pages users are likely to view. In particular, the "people also like" feature may guide users to certain pages. Given that it lists similar types of pages as well (other non-profit organizations in the example above), it may also help users find certain types of pages. Future research will be required to examine how changes in search engines impact users' ability to find relevant information and pages.

Second, we evaluated the representation of health conditions only on Facebook pages and we did not examine private messages or private groups as we were interested in what is made public to all users when searching for health conditions on Facebook. Private messaging and groups might be more appropriate channels for communicating about sensitive health topics and warrant future research. This limitation is not unique to our study; previous research on health conditions among Facebook groups has focused on public groups and messages. Additionally, our data were collected over a limited time period and for only the first 50 pages returned in the search results. A more comprehensive set of data may yield evidence of longitudinal or seasonal patterns in the representation of health conditions that we were not able to detect.

Third, we did not attempt to formally evaluate the accuracy of the informational content of pages. Thus, pages that were categorized as relevant may vary greatly in the utility of the information provided for differing health conditions. Future research on the quality of content across key health conditions may highlight critical topics of misinformation and be used to support interventions designed to correct and/or counter poor information resources.

Fourth, our descriptive analysis does not provide any data on the characteristics of Facebook members who searched for health conditions or how they used the information they found. Future research should examine how people make use of Facebook as one element of a communication ecology to address their informational needs and to garner social support, and how this usage impacts their health care utilization, self-care, and health outcomes.

Conclusions

The rapid growth and diffusion of social media and SNS during the past 10 years has created new opportunities for people to find and share information about a wide variety of health conditions. Facebook is the most widely used SNS in the United States; however, little is known about the diversity of health conditions represented on Facebook. This research represents the first attempt to comprehensively describe the content and level of user engagement with health conditions on Facebook pages. Our findings provide useful baseline information and several insights to inform future research and interventions designed to improve public health. We found that a search of Facebook for common health conditions provided a large number of irrelevant pages. In addition, most pages were devoted to marketing/promotion and relatively few pages were devoted to social support. Social support was especially underrepresented in pages for health conditions for communicable diseases. Public health interventions using Facebook will need to be designed to ensure relevant information is easy to find and with an understanding that stigma associated with some health conditions may limit the utility of Facebook as a channel for health communication. This line of research merits further investigation as Facebook and other SNS continue to evolve over the coming years.

Conflicts of Interest

None declared.

References

1. Fox S, Duggan M. Pew Internet and American Life Project. Washington, DC: Pew Research Center; 2013. Health online 2013 URL: http://www.pewinternet.org/~/media//Files/Reports/PIP_HealthOnline.pdf [accessed 2014-01-08] [WebCite Cache ID 6MTHCH4MT]
2. Hesse BW, Moser RP, Rutten LJ. Surveys of physicians and electronic health information. N Engl J Med 2010 Mar 4;362(9):859-860. [doi: 10.1056/NEJMc0909595] [Medline: 20200398]
3. Cotten SR. Implications of internet technology for medical sociology in the new millennium. Sociological Spectrum 2001 Jul;21(3):319-340. [doi: 10.1080/027321701300202019]
4. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
5. Brenner J. Pew Internet and American Life Project. Washington, DC: Pew Research Center; 2013. Social networking 2013 URL: http://pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx [accessed 2014-01-08] [WebCite Cache ID 6MTPLgHrQ]
6. Duggan M, Brenner J. Pew Internet and American Life Project. Washington, DC: Pew Research Center The demographics of social media users 2012 URL: http://pewinternet.org/~/media//Files/Reports/2013/PIP_SocialMediaUsers.pdf [accessed 2014-01-08] [WebCite Cache ID 6MTPFfEKi]
7. Facebook.: Facebook, Inc; 2014. Key Facts - Facebook's latest news, announcements and media resources URL: http://newsroom.fb.com/Key-Facts [accessed 2014-01-08] [WebCite Cache ID 6MTPYAhzE]
8. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]

9. Ahmed OH, Sullivan SJ, Schneiders AG, McCrory P. iSupport: do social networking sites have a role to play in concussion awareness? Disabil Rehabil 2010 Jan;32(22):1877-1883. [doi: 10.3109/09638281003734409] [Medline: 20367328]
10. Gajaria A, Yeung E, Goodale T, Charach A. Beliefs about attention-deficit/hyperactivity disorder and response to stereotypes: youth postings in Facebook groups. J Adolesc Health 2011 Jul;49(1):15-20. [doi: 10.1016/j.jadohealth.2010.09.004] [Medline: 21700151]
11. Bender JL, Jimenez-Marroquin MC, Jadad AR. Seeking support on Facebook: a content analysis of breast cancer groups. J Med Internet Res 2011;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]
12. Thoren EM, Metze B, Bührer C, Garten L. Online support for parents of preterm infants: a qualitative and content analysis of Facebook 'preemie' groups. Arch Dis Child Fetal Neonatal Ed 2013 Nov;98(6):F534-F538. [doi: 10.1136/archdischild-2012-303572] [Medline: 23759520]
13. De la Torre-Díez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]
14. Farmer AD, Bruckner Holt CE, Cook MJ, Hearing SD. Social networking sites: a novel portal for communication. Postgrad Med J 2009 Sep;85(1007):455-459. [doi: 10.1136/pgmj.2008.074674] [Medline: 19734511]
15. Frankhauser D. Mashable.com. 2013. 18 Facebook fossils we'll remember forever URL: http://mashable.com/2013/01/25/facebook-fossils/ [accessed 2014-01-08] [WebCite Cache ID 6MTPgMrJF]
16. Warble A. Notes by Facebook Engineering.: Facebook, Inc; 2010 Mar 16. Notes by Facebook engineering: intro to Facebook search URL: http://www.facebook.com/notes/facebook-engineering/intro-to-facebook-search [accessed 2014-05-11] [WebCite Cache ID 6PUpifjMc]
17. Sajadi KP, Goldman HB. Social networks lack useful content for incontinence. Urology 2011 Oct;78(4):764-767. [doi: 10.1016/j.urology.2011.04.074] [Medline: 21862115]
18. Dobransky K, Hargittai E. Inquiring minds acquiring wellness: uses of online and offline sources for health information. Health Commun 2012;27(4):331-343. [doi: 10.1080/10410236.2011.585451] [Medline: 21932982]
19. Percheski C, Hargittai E. Health information-seeking in the digital age. J Am Coll Health 2011;59(5):379-386. [doi: 10.1080/07448481.2010.513406] [Medline: 21500056]
20. Robinson L. A taste for the necessary: a Bourdieuian approach to digital inequality. Information, Communication & Society 2009 Jun;12(4):488-507. [doi: 10.1080/13691180902857678]
21. Hargittai E, Hinnant A. Digital inequality: differences in young adults' use of the Internet. Communication Research 2008 Aug 04;35(5):602-621. [doi: 10.1177/0093650208321782]
22. DiMaggio P, Hargittai E, Celeste C, Shafer S. Digital inequality: from unequal access to differentiated use. In: Neckerman KM, editor. Social Inequality. New York, NY: Russell Sage Foundation; 2004:355-400.
23. Viswanath K. Public communications and its role in reducing and eliminating health disparities. In: Thompson GE, Michell F, Williams MB, editors. Examining the health disparities research plan of the National Institutes of Health: unfinished business. Washington, DC: National Academy Press; 2006:215-252.
24. Viswanath K. Science and society: the communications revolution and cancer control. Nat Rev Cancer 2005 Oct;5(10):828-835. [doi: 10.1038/nrc1718] [Medline: 16195753]
25. Rains SA. The implications of stigma and anonymity for self-disclosure in health blogs. Health Commun 2014;29(1):23-31. [doi: 10.1080/10410236.2012.714861] [Medline: 23356432]
26. Lampe C, Vitak J, Gray R, Ellison N. Perceptions of Facebook's value as an information source. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.: ACM; 2012 Presented at: CHI; May 5-10, 2012; Austin, TX p. 3195-3204. [doi: 10.1145/2207676.2208739]

Abbreviations

ADHD: attention deficit hyperactivity disorder
AIDS: acquired immunodeficiency syndrome
HIV: human immunodeficiency virus
HPV: human papillomavirus
IRR: interrater reliability
SNS: social networking sites

Edited by G Eysenbach; submitted 25.01.14; peer-reviewed by WYS Chou, J Loss, I de la Torre; comments to author 28.04.14; revised version received 31.05.14; accepted 14.07.14; published 04.08.14.

Please cite as:
Hale TM, Pathipati AS, Zan S, Jethwani K
Representation of Health Conditions on Facebook: Content Analysis and Evaluation of User Engagement
J Med Internet Res 2014;16(8):e182
URL: http://www.jmir.org/2014/8/e182/
doi:10.2196/jmir.3275
PMID:25092386

Original Paper

Changes Over Time in the Utilization of Disease-Related Internet Information in Newly Diagnosed Breast Cancer Patients 2007 to 2013

Christoph Kowalski[1], PhD; Eva Kahana[2], PhD; Kathrin Kuhr[3], MSc; Lena Ansmann[1], PhD; Holger Pfaff[1], PhD

[1]IMVR, University of Cologne, Koeln, Germany

[2]Department of Sociology and Elderly Care Research Center, Case Western Reserve University, Cleveland, OH, United States

[3]Institute of Medical Statistics, Informatics and Epidemiology, University of Cologne, Koeln, Germany

Corresponding Author:
Christoph Kowalski, PhD
IMVR
University of Cologne
Eupener Str 129
Koeln, 50933
Germany
Phone: 49 221 47897148
Fax: 49 221 47897142
Email: christoph.kowalski@uk-koeln.de

Abstract

Background: As the number of people with Internet access rises, so does the use of the Internet as a potentially valuable source for health information. Insight into patient use of this information and its correlates over time may reveal changes in the digital divide based on patient age and education. Existing research has focused on patient characteristics that predict Internet information use and research on treatment context is rare.

Objective: This study aims to (1) present data on the proportion of newly diagnosed breast cancer patients treated in German breast centers from 2007 to 2013 who used the Internet for information on their disease, (2) look into correlations between Internet utilization and sociodemographic characteristics and if these change over time, and (3) determine if use of Internet information varies with the hospitals in which the patients were initially treated.

Methods: Data about utilization of the Internet for breast cancer–specific health information was obtained in a postal survey of breast cancer patients that is conducted annually in Germany with a steady response rate of 87% of consenting patients. Data from the survey were combined with data obtained by hospital personnel (eg, cancer stage and type of surgery). Data from 27,491 patients from 7 consecutive annual surveys were analyzed for this paper using multilevel regression modeling to account for clustering of patients in specific hospitals.

Results: Breast cancer patients seeking disease-specific information on the Internet increased significantly from 26.96% (853/3164) in 2007 to 37.21% (1485/3991) in 2013. Similar patterns of demographic correlates were found for all 7 cohorts. Older patients (≥70 years) and patients with <10 years of formal education were less likely to use the Internet for information on topics related to their disease. Internet use was significantly higher among privately insured patients and patients living with a partner. Higher cancer stage and a foreign native language were associated with decreased use in the overall model. Type of surgery was not found to be associated with Internet use in the multivariable models. Intraclass correlation coefficients were small (0.00-0.03) suggesting only a small contribution of the hospital to the patients’ decision to use Internet information. There was no clear indication of a decreased digital divide based on age and education.

Conclusions: Use of the Internet for health information is on the rise among breast cancer patients. The strong age- and education-related differences raise the question of how relevant information can be adequately provided to all patients, especially to those with limited education, older age, and living without a partner.

(J Med Internet Res 2014;16(8):e195) doi:10.2196/jmir.3289

KEYWORDS

disease-related Internet information; digital divide; multilevel analysis; breast neoplasms; hospital differences

Introduction

As the number of people with Internet access continues to rise, so too does the number of people using the Internet as a source for health information or health-related activities [1,2]. Over the past two decades, this has led to major changes in both the way health information is consumed and the amount of knowledge laypersons can access relatively easily [3,4].

Breast cancer offers an important arena for exploration of patient Internet use. Breast cancer is a major public health concern, as it is the most common form of cancer and the second major cause of cancer-related deaths among women in the United States. In Germany, 1 in 8 women will face a breast cancer diagnosis in her lifetime [5]. In the Internet age, a new role has become available to patients as information managers. Information acquisition through the Internet can help develop patient competence in dealing with challenges of a life-threatening illness, such as breast cancer [6].

Internet accessibility and its use for health purposes are distributed unequally over the population and its effects are not without controversies. Focusing on benefits to patients, a number of studies emphasize that using health information from the Internet is associated with stronger participation in decision making [7], better decisions [8], more frequent change of health behavior [9], and it may enable patients to communicate with doctors more effectively [10,11]. In contrast, other studies argue that using the Internet for health information may lead to erosion of the patient-provider relationship [12,13] or may confuse patients [14]. The early literature on health-related Internet use was particularly concerned with the limited ability of laypersons to evaluate information obtained on the Internet [15].

It is increasingly important for health care providers to give serious consideration to the information patients collect and to address their understanding of that information [16-18]. Taking into account the varying quality of websites providing health information, quality assurance and expert participation is warranted [19]. Nevertheless, there is indication of improvement in the quality of information offered to patients with breast cancer through a growing number of high quality websites (eg, National Institutes of Health [20], Agency for Healthcare Research and Quality [21], and National Cancer Institute [22] in the United States, and gesundheitsinformation.de [23] and Krebsgesellschaft [24] in Germany). As Eysenbach stated, referring to the accuracy of cancer information websites as far back as 2003: it “is not so bad after all” [10]. The increasing sophistication of Internet sites enables patients to access not only sites designed for patients, but also peer-reviewed scientific articles that describe the latest research relevant to specific problems of the patient.

Patients using the Internet to gain access to health information for various illnesses tend to be younger and of higher socioeconomic status across countries [25-31]. This well-documented “digital divide” might become a major threat to equity in health care once relevant or even necessary information can only be or best be accessed online.

Although reports on the proportion of patients who use the Internet to gain health information vary widely [10,32,33], recent results based on 2011 data suggest that more than 50% of breast cancer patients [25] used the Internet to gain disease-specific information. Because of such widespread reliance on the Internet among female breast cancer patients, there is a clear need for up-to-date information on trends in this form of information acquisition. Differences in the proportion of individuals using the Internet not only differ according to the specific sample and the country or region under investigation, but also study design and the questions posed. Variability among studies in the nature of the disease and time since diagnosis also makes comparisons over time difficult and leaves unanswered questions about trends in the digital divide in relation to health information-seeking [34,35]. Although much research focuses on demographic correlates of online health information use, to our knowledge no study has yet investigated differences across locations of treatment. If variation across locations of treatment persists after controlling for individual characteristics this might offer further important clues to patient motivations for using the Internet for information. Thus, it is possible that unsatisfying experiences in the medical encounter or limited explanations communicated by health care providers would result in increased patient Internet use for health-related information.

The aim of our study was to expand the knowledge base about personal demographic, contextual, and temporal determinants of Internet use among newly diagnosed cancer patients. Specifically, this study aims to (1) present data on the proportion of 7 cohorts of newly diagnosed breast cancer patients treated in German breast center hospitals from 2007 to 2013 who used the Internet for information on their disease, (2) consider stability and change in patient characteristics predicting Internet use over time focusing on the digital divide based on age, education, and insurance status as an indicator of socioeconomic status, and (3) determine if use of information from the Internet varies by the hospital in which the patients were initially treated.

In doing so, we hope to expand existing knowledge by investigating developments over time and addressing the health care organization’s contribution to online health information use while taking clinical data (stage, type of surgery) and potentially relevant patient characteristics (partnership status, native language, gender) into account.

Methods

Participants

This report analyzed data drawn from a larger program of research designed to investigate the breast center concept of the German federal state of North Rhine-Westphalia (population 17.5 million). Patients treated for newly diagnosed breast cancer in one of the accredited breast center hospitals were asked to self-administer a questionnaire at home after discharge from the hospital [36]. Patients were included in the survey if they had a first diagnosis of breast cancer, underwent surgery during their current hospital stay, and had at least one malignancy, at least one postoperative histology, and a confirmed diagnosis of breast cancer with an International Classification of Diseases (ICD) code of C50.x or D05.x. Each year between February and June (survey period 6 months), all patients who fulfilled the inclusion criteria were included in the study consecutively. Cross-sectional surveys were performed with samples of patients from all accredited breast center hospitals in the region studied.

Shortly before discharge from the hospital, patients were asked by the hospital staff to give written consent to be included in the survey. Once the patients had given their consent, hospital personnel from the centers provided the research team with clinical information on the patients. The survey was designed according to Dillman's Total Design Method with 3 contacts [37]. The survey was sent out to the patient's home address within a week of receiving written consent. The study was approved by the institutional ethics committee of the University Hospital of Cologne, Germany. We analyzed data from each of the 7 years (2007 to 2013). Of the 35,371 patients meeting the inclusion criteria, 31,293 (88.47%) consented to the survey. Of these, 27,491 (87.85%) returned the questionnaire. These patients make up the sample for the analyses.

Measures

Dependent Variable

The dependent variable was use of the Internet for breast cancer–specific health information assessed based on response to a survey question that asked about such Internet use (yes/no).

Independent Variables

Patient sociodemographic data and clinical status served as independent variables. Patients were asked to indicate their date of birth, native language, insurance status, highest year of education attained, and partnership status on the questionnaire. Except for age (continuous) the sociodemographic variables are categorized into native language (German vs other), insurance status (statutory health insurance vs partly private/partly private), highest year of education (≥10 years of school vs <10 years of school), partnership status (living with a partner vs not living with a partner), and gender (male vs female).

In addition to the data collected by the patient questionnaire, medical personnel contributed clinical data and information about type of surgery performed after patient consent. The cancer stage was categorized using Union for International Cancer Control (UICC) categories [38]; type of surgery was dichotomized (breast-conserving treatment vs mastectomy).

Statistical Analyses

Proportions of Internet Users

The proportion of patients who used Internet information about breast cancer was calculated separately for each of the 7 cohorts, both overall and stratified for younger patients with more formal education to spotlight the digital divide (age <50 years; ≥10 years of school) and older patients with less formal education (age ≥70 years; <10 years of school). To test for differences over time, the Cochran-Armitage trend test was applied. In addition, the share of the 4 groups that resulted when stratifying for age and education among Internet users was analyzed. We performed bivariate tests to examine associations between the independent variables included in the model. We conducted chi-square tests for associations between all categorical variables (type of surgery, native language, years of schooling, insurance status, living with a partner, gender, cancer stage). Spearman rank correlation was used to examine the correlation between age and the ordinal variable cancer stage. Also, *t* tests were conducted to examine age differences for the different groups in the dichotomous variables. The cross-year dataset was used for these analyses.

Multilevel Models

Data from each survey cohort were analyzed separately and in an overall model using multilevel analysis. This is the method of choice when accounting for the nested structure of the data, such as patients (level 1) in hospitals (level 2) [39]. Two-level models without predictors were fitted to yield the intraclass correlation coefficient (ICC) for the null model. The ICC represents the proportion of the variance of the dependent variable attributable to the hospital level. In a second step, all patient characteristics were included. A number of patients indicated they did not have access to the Internet in an earlier question and did not respond to the dependent variable. To avoid case deletion, cases that indicated they did not have access to the Internet in the earlier question were coded as not having used the Internet. Cases with missing data in the dependent variable and missing data in this earlier question were excluded from all analyses (n=1022). Patients with missing data in the continuous age variable were excluded in the multilevel models (n=243), leaving 26,226 patients for the multilevel analyses. Missing data on all other independent variables were included in the model as separate categories to avoid case deletion, and omitted in the results tables. The ICCs of these models represent the proportion of variance attributable to the hospital-level characteristics after accounting for variation in the patient characteristics, (ie, the different patient case mix). Because of the small ICCs, no hospital-level characteristics were included in the models. The overall model included a cohort variable to account for the survey year. In addition, we included a gender variable that we did not include in the year-by-year analyses because of small strata. SPSS version 22.0 (IBM Corp, Armonk, NY, USA) was used for descriptive analysis and MLWiN 2.25 (Centre for Multilevel Modelling, Bristol, UK) for multilevel analysis. R 3.0.2 (R Project for Statistical Computing, Vienna, Austria) was used to calculate the Cochran-Armitage trend test.

Results

Table 1 shows the percentage of breast cancer patients who reported they used the Internet to obtain information about their disease. There was a relatively steady, statistically significant increase in this percentage over the 7-year study period (2007: 26.96%, 853/3164; 2013: 37.21%, 1485/3991; χ^2_1=138.0, *P*<.001). No relevant changes were found for the proportion of younger, higher-educated patients who used the Internet (χ^2_1=0.4, *P*=.51). Proportions for this group remained relatively stable, between 60% and 70% throughout the study period. The proportion of older patients with little formal education who used Internet information increased significantly from 2007 (2.9%, 13/444) to 2013 (4.7%, 29/617; χ^2_1=6.8, *P*=.009) but remained below 6% for all cohorts. Among men, the overall proportion was only 25.4% (32/126, not presented in a table).

Table 1. Patients reporting to have used the Internet to obtain information about breast cancer across the entire study period (2007-2013) and by younger, higher-educated patients and older, less-educated patients.

Year	Overall		Age <50 years; ≥10 years of school		Age ≥70 years; <10 years of school	
	n/N	%	n/N	%	n/N	%
2007	853/3164	26.96	283/452	62.6	13/444	2.9
2008	1093/3689	29.63	331/547	60.5	14/517	2.7
2009	1196/3855	31.02	355/559	63.5	20/614	3.3
2010	1272/3767	33.77	407/606	67.2	30/576	5.2
2011	1343/3940	34.09	397/635	62.5	26/668	3.9
2012	1505/4063	37.04	413/634	65.1	37/666	5.6
2013	1485/3991	37.21	401/642	62.5	29/617	4.7

To better understand which patient group contributed most to the increase in Internet use, we compared the share in users for 4 different groups: (1) age ≥70 years, <10 years of school; (2) age <50 years, <10 years of school; (3) age ≥70 years, ≥10 years of school; and (4) age <50 years, ≥10 years of school (Table 2). None of the 4 groups' share of Internet users increased substantially over time.

Table 2. Composition of Internet health information users.

Patient subgroup[a]	Year, n/N[b] (%)						
	2007	2008	2009	2010	2011	2012	2013
Age ≥70 years; <10 years of school	13/824 (1.6)	14/1074 (1.30)	20/1181 (1.69)	30/1268 (2.37)	26/1333 (1.95)	37/1498 (2.47)	29/1464 (1.98)
Age <50 years; <10 years of school	54/824 (6.6)	74/1074 (6.89)	74/1181 (6.27)	75/1268 (5.91)	72/1333 (5.40)	63/1498 (4.21)	63/1464 (4.30)
Age ≥70 years; ≥10 years of school	10/824 (1.2)	20/1074 (1.86)	20/1181 (1.69)	26/1268 (2.05)	41/1333 (3.08)	60/1498 (4.01)	68/1464 (4.64)
Age <50 years; ≥10 years of school	283/824 (34.3)	331/1074 (30.82)	355/1181 (30.06)	407/1268 (32.10)	397/1333 (29.78)	413/1498 (27.57)	401/1464 (27.39)

[a] Patients aged 50 to 69 years comprise the remaining portion of the sample.

[b] The N's presented represent only the patients with valid data for education and age.

Tables 3 to 5 present bivariate associations between the independent variables in the sample. Most notably, partnership and insurance status were significantly correlated with many other study variables, such as type of surgery, native language, stage, and education. Age differences were found for type of surgery, native language, years of schooling, partnership status, and gender. Spearman rho was .069 (*P*<.001) for the correlation between stage (ordinal) and age (not presented in a table).

Table 3. Bivariate associations between independent variables in the sample: dichotomous variables, percentages, and Pearson chi-square[a] (N=27,491).

Variable	Mastectomy, %				Native language, %				Years of school, %				Private health insurance,[c] %				Living with partner, %			
	Yes	No[b]	χ^2_1	*P*	German	Other	χ^2_1	*P*	<10	≥10	χ^2_1	*P*	Yes	No	χ^2_1	*P*	No	Yes	χ^2_1	*P*
Gender																				
Female	26.6	73.4	285.9	<.001	94.6	5.4	3.5	0.08	51.6	48.4	0.7	0.43	24.6	75.4	2.0	0.12	28.1	71.9	7.6	.006
Male	93.7	6.3			90.8	9.2			55.4	44.6			29.9	70.1			17.3	82.7		
Mastectomy																				
Yes					95.0	5.0		0.09	52.9	47.1	5.2	0.02	23.2	76.8	10.8	0.001	33.9	66.1		<.001
No[b]					94.5	5.5	3.0		51.3	48.7			25.2	74.8			26.0	74.0	155.4	
Native language																				
German									52.4	47.6	121.8	<.001	25.5	74.5	179.8	<.001	28.3	71.7	26.0	<.001
Other									37.1	62.9			9.8	90.2			22.2	77.8		
Years of school																				
<10													13.7	86.3	1883.8	<.001	29.4	70.6	33.4	<.001
≥10													37.1	62.9			26.2	73.8		
Private health insurance [c]																				
Yes																	24.4	75.6	58.7	<.001
No																	29.4	70.6		

[a] Pairwise deletion used in the chi-square analysis, Fisher's exact test.

[b] Breast-conserving treatment.

[c] Yes: (partly) private health insurance; no: only statutory health insurance.

Table 4. Bivariate associations between cancer stage and dichotomous independent variables in the sample: percentages and Pearson chi-square[a] (N=27,491).

Variable	Cancer stage, n/N (%)					χ^2_4	*P*
	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4		
Mastectomy						3078.1	<.001
Yes	426/1484 (28.71)	1264/9953 (12.70)	2394/8029 (29.82)	1541/2505 (61.52)	496/806 (61.54)		
No[b]	1058/1484 (71.29)	8689/9953 (87.30)	5635/8029 (70.18)	964/2505 (38.48)	310/806 (38.46)		
Native language						7.2	
German	1405/1480 (94.93)	9434/9930 (95.01)	7624/8096 (94.17)	2386/2529 (94.35)	798/839 (95.11)		
Other	75/1480 (5.07)	496/9930 (4.99)	472/8096 (5.83)	143/2529 (5.65)	41/839 (4.89)		
Years of school						53.9	<.001
<10	701/1473 (47.59)	5001/9878 (50.63)	4145/7936 (52.23)	1386/2480 (55.89)	493/825 (59.76)		
≥10	772/1473 (52.41)	4877/9878 (49.37)	3791/7936 (47.77)	1094/2480 (44.11)	332/825 (40.24)		
Private health insurance [c]						55.0	<.001
Yes	390/1458 (26.75)	2570/9712 (26.46)	1828/7912 (23.10)	545/2481 (21.97)	158/834 (18.94)		
No	1068/1458 (73.25)	7142/9712 (73.54)	6084/7912 (76.90)	1936/2481 (78.03)	676/834 (81.06)		
Living with partner						88.4	<.001
No	380/1478 (25.71)	2515/9947 (25.28)	2374/8062 (29.45)	837/2515 (33.28)	261/830 (31.45)		
Yes	1098/1478 (74.29)	7432/9947 (74.72)	5688/8062 (70.55)	1678/2515 (66.72)	569/830 (68.55)		
Gender						32.9	<.001
Male	4/1510 (0.26)	31/10,127 (0.31)	43/8252 (0.52)	27/2580 (1.05)	10/861 (1.2)		
Female	1506/1510 (99.74)	10,096/10,127 (99.69)	8209/8252 (99.48)	2553/2580 (98.95)	851/861 (98.8)		

[a] Pairwise deletion used in the chi-square analysis.

[b] Breast-conserving treatment.

[c] Yes: (partly) private health insurance; no: only statutory health insurance.

Table 5. Bivariate associations between age and dichotomous independent variables in the sample (N=27,491).

Variable	Mean age	*t* (*df*)	*P*
Mastectomy			<.001
Yes	62.5	–14.96 (10,537.5)	
No[a]	59.7		
Native language			<.001
German	60.7	–20.88 (1666.8)	
Other	54.6		
Years of school			<.001
<10	64.5	–63.59 (25,574.3)	
≥10	55.8		
Private health insurance [b]			.587
Yes	60.5	–0.54 (11,474.2)	
No	60.4		
Living with partner			<.001
No	64.6	–35.83 (11,961.4)	
Yes	58.6		
Gender			<.001
Male	66.2	5.62 (27,216)	
Female	60.4		

[a] Breast-conserving treatment.

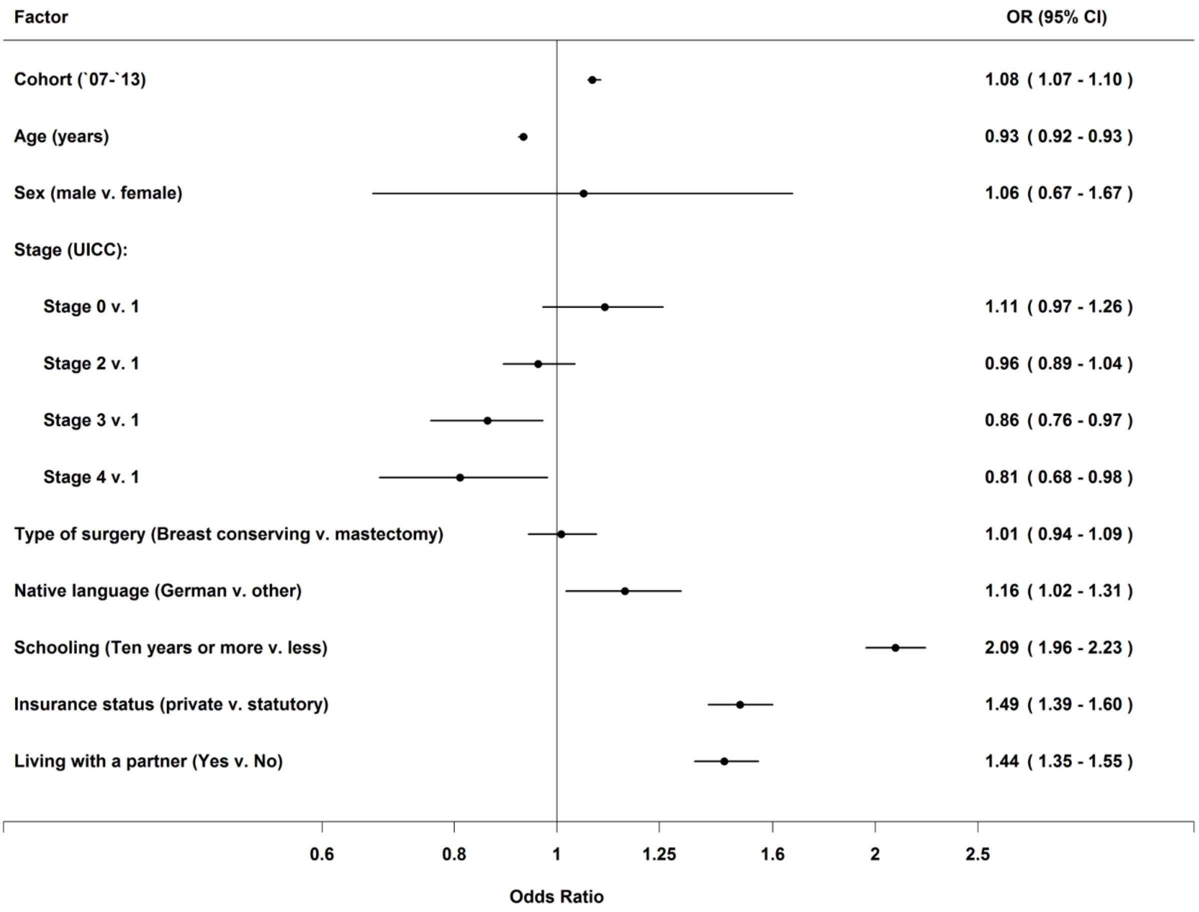
[b] Yes: (partly) private health insurance; no: only statutory health insurance.

Results from the multilevel analyses are presented in Table 6 and Figure 1. Table 6 presents results for each single cohort and Table 1 for the overall model (ie, an across-years analysis that includes an additional year variable). Table 6 and Figure 1 reveal associations between health-related Internet use and age and education, with higher formal education (OR 2.09, 95% CI 1.96-2.23) and decreasing age (OR 0.93, 95% CI 0.92-0.93) being significantly associated with higher Internet information use in the overall model. In addition, patients who were privately or partly privately insured (OR 1.49, 95% CI 1.39-1.60) or were living with a partner (OR 1.44, 95% CI 1.35-1.55) were more likely to use the Internet for breast cancer–related information. Each of the cohorts yielded the same statistically significant predictors of Internet use (except for stage in 2011 and partner in 2013) with only slight differences in effect sizes. A foreign native language vs German (OR 1.16, 95% CI 1.02-1.31) and cancer stages 3 (OR 0.86, 95% CI 0.76-0.97) and 4 (OR 0.81, 95% CI 0.68-0.98) vs cancer stage 1 were found to be statistically significantly associated with less Internet use in the overall model only. Type of surgery (OR 1.01, 95% CI 0.94-1.09) and gender (OR 1.06, 95% CI 0.67-1.67) were not associated with Internet use. ICCs of the dependent variable in the null models were small (≤0.03) for all 7 cohorts, especially after introducing patient level predictors, indicating small differences between hospitals (2007: 0.03 after including patient characteristics, 0.03 for the null model; 2008: 0.01, 0.03; 2009: 0.01, 0.02; 2010: 0.00, 0.01; 2011: 0.00, 0.01; 2012: 0.02, 0.03; 2013: 0.01, 0.03).

Table 6. Logistic multilevel regression analyses on having used the Internet to obtain information about the breast cancer (N=26,226).

Variable	Year, OR (95% CI)						
	2007	2008	2009	2010	2011	2012	2013
	n=3078	n=3630	n=3816	n=3760	n=3920	n=4053	n=3969
Constant							
	15.27 (7.20-32.41)	23.82 (11.98-47.35)	24.00 (12.35-46.68)	15.37 (8.15-28.98)	11.62 (6.33-21.35)	13.70 (7.56-24.83)	21.04 (11.67-37.96)
Age (years)							
	0.92 (0.91-0.93)	0.92 (0.91-0.93)	0.93 (0.92-0.93)	0.93 (0.92-0.94)	0.93 (0.93-0.94)	0.94 (0.93-0.94)	0.93 (0.93-0.94)
Stage (ref: stage 1)							
Stage 0	0.94 (0.61-1.43)	1.04 (0.74-1.45)	1.37 (0.97-1.94)	0.95 (0.64-1.41)	1.36 (0.97-1.92)	1.06 (0.78-1.45)	1.10 (0.81-1.51)
Stage 2	1.15 (0.91-1.44)	0.94 (0.77-1.15)	0.95 (0.78-1.16)	1.07 (0.88-1.30)	0.94 (0.78-1.13)	0.86 (0.71-1.03)	0.95 (0.79-1.44)
Stage 3	0.90 (0.63-1.27)	0.78 (0.57-1.07)	0.81 (0.59-1.12)	0.82 (0.60-1.12)	0.97 (0.72-1.31)	0.85 (0.63-1.14)	0.80 (0.59-1.08)
Stage 4	1.60 (0.89-2.86)	0.64 (0.37-1.13)	1.19 (0.72-1.97)	1.02 (0.66-1.58)	0.60 (0.37-0.97)	0.67 (0.39-1.13)	0.73 (0.46-1.15)
Type of surgery							
Breast conserving (vs mastectomy)	0.90 (0.72-1.12)	1.20 (0.97-1.48)	0.96 (0.79-1.17)	1.05 (0.86-1.28)	0.93 (0.77-1.13)	0.98 (0.82-1.18)	1.00 (0.83-1.21)
Native language							
German (vs other)	1.32 (0.87-2.02)	0.96 (0.67-1.38)	1.02 (0.71-1.45)	1.30 (0.93-1.81)	1.34 (0.98-1.83)	1.02 (0.76-1.37)	0.91 (0.69-1.21)
Years of schooling							
≥10 (vs <10)	2.40 (1.96-2.95)	1.92 (1.60-2.30)	2.13 (1.79-2.53)	1.84 (1.55-2.17)	1.99 (1.68-2.35)	2.02 (1.71-2.38)	2.20 (1.86-2.59)
Insurance status							
(Partly) private (vs statutory)	1.31 (1.06-1.62)	1.83 (1.50-2.22)	1.31 (1.07-1.58)	1.84 (1.53-2.21)	1.37 (1.14-1.65)	1.28 (1.07-1.53)	1.48 (1.25-1.76)
Living with a partner							
Yes (vs no)	1.50 (1.19-1.88)	1.44 (1.18-1.76)	1.38 (1.14-1.68)	1.31 (1.09-1.57)	1.47 (1.23-1.77)	1.52 (1.27-1.81)	1.13 (0.96-1.34)

Figure 1. Logistic multilevel regression analysis on having used the Internet to obtain information about the disease for the overall model cohorts from 2007-2013 (N=26,226).



Discussion

Findings of this study expand the discussion about the role of computer technology to facilitate proactive illness management for breast cancer patients. Patients diagnosed with breast cancer have creatively used the Web to meet the challenges of their illness. It has been argued that these patients face substantial barriers as they try to make sense of their illness in a fragmented and limited information environment [40].

Our findings confirm expectations that the use of the Internet for seeking health information is on the rise among breast cancer patients. As hypothesized, being younger and having a higher level of education increased the likelihood that a patient would search the Internet for information about their disease. Two-thirds of patients younger than 50 years and with more than 10 years of schooling used the Internet, whereas less than 6% of those older than 70 years and with less than 10 years of schooling did so in all 7 cohorts, with only small changes over time. In addition, living with a partner and having private insurance was positively associated with Internet information use. Only small variation in the dependent variable was found between hospitals and over time. This indicates that there is no systematic impact of the treating institution on the patients' decision to use Internet information. In aggregate, these data indicate that personal demographic factors play a much greater role in shaping proactive involvement in searching for health information than do situational aspects of the health care environment.

We were able to show a small association for cancer stage and Internet use only in the overall model, pointing to a more limited role of illness characteristics among personal determinants of Internet use [30]. Studies with population-based samples often found that individuals who reported impaired health or chronic conditions used the Internet more frequently (eg, [41,42]). Our sample consisted of individuals who suffered from an acute and life-threatening disease; therefore, they were relatively homogenous with regard to the health status.

A substantial digital divide was found in our study with respect to age and formal education and it did not clearly decrease over time. This finding is inconsistent with suggestions in the literature that the digital divide may be disappearing [43]. Given that partnership status of the patients in our study contributed to Internet use, it is apparent that some patient groups are systematically excluded from one of the most common contemporary sources of information. This raises concerns about alternative methods for meeting pressing information needs of patients encountering a diagnosis that poses great uncertainty. Furthermore, these patients have little access to interventions and practical tools involving computers or advanced electronic devices, such as smartphones or tablet computers [44]. Discovering how these patient groups can be adequately approached is a central task for future health communication efforts.

Privately insured patients used the Internet to search for information on breast cancer significantly more often than their counterparts. This finding is consistent with prior research [45] and is most likely due to the higher socioeconomic status of privately insured patients. This is not fully captured by the education variable in the models. A population-based survey from the United States showed that people living in rural areas used the Internet less than their urban counterparts did both in general and for health-specific purposes. This difference is mainly due to the differences in socioeconomic status and accessibility of broadband [46]. This not only jeopardizes aims of equal access to health information, but might also lead to worse quality of care and a confounding factor when comparing providers. Our data do not suggest that such an effect exists in our specific sample, since differences between hospitals (which controls for rural/urban differences to some extent) were small.

Patients who were not native speakers of German were found to use the Internet less often only in the overall model. However, it must be considered that patients with difficulties understanding German are likely to be underrepresented in this sample because the questionnaire was administered in German. This is relevant to interpreting bivariate associations with the native language variable. However, statistically adjusting for this would require more knowledge about nonrespondents with a native language other than German. Research has shown that the degree to which a person is comfortable speaking a language other than his or her own affects both use of and trust in health information sources [47]. This does not contradict the previously tested diversification hypothesis in which minority status, not native language, proven to shape health-relevant Internet use [48]. The finding of lesser Internet use in nonnative speakers as well as the strong decrease of Internet information use with decreasing formal education also reflects inaccessibility of Internet based health information to patients with low health literacy. Much of the information that is accessible online exceeds the reading level that is recommended for general use by information-seeking patients [49,50].

It has been reported that the prevalence of seeking health information on the Internet is higher in women compared to men [51]. In our sample, we found no gender differences when taking other patient characteristics into account. Because male breast cancer patients tend to have difficulties accessing important information through traditional channels [52,53], it is somewhat surprising that they do not use the Internet more than female patients.

Online information can be a central resource for the elderly who may have difficulty in accessing health information because they are homebound and/or have little social support [54]. This requires a careful investigation of what might help increase the number of older adult Internet users. Some attention has been given to factors that contribute to the lower rates of Internet information use among the older population besides physical impairment, less access, and less familiarity. For example, older adults may distrust the information provided online [55]. Select, expert-guided, quality assured information that is recommended by health care providers might be a key to reaching this skeptical patient subgroup. Additionally, older adults may benefit from training and from availability of more senior-friendly design. This strategy has long been advocated [56].

The third group that needs to be focused on is the part of the population with low formal education and limited health literacy to avoid the reinforcement of existing social differences [57,58]. The more information that is available online, the more important it is to also provide them through other sources of communication for those who do not have access to the Internet or are not Internet savvy.

A number of limitations of our study need to be mentioned. The study sample is a highly specific subsample of the general population and this limits generalizability of findings. Another limitation is the general nature of the Internet usage measure that does not specify types of information that was actually accessed and how this relates to patient preferences [59]. As each cohort was surveyed only once in our study, we cannot establish whether seeking Internet information is a result of the experience within the hospital or independent of it. Further research is required to investigate the temporal order [18,60]. However, the small ICCs hint at little impact of the institutional context. Also, comparing proportions of patients using the Internet to obtain disease-specific information with results from other studies should be done with caution. Patients in this study responded to the survey shortly after discharge from the hospital and it is possible that some patients consulted the Internet later on. We were able to detect significant associations between Internet use and native language as well as cancer stage only in the overall model with the higher statistical power as compared to the year-by-year analyses.

Despite its limitations, the present study is ground breaking in providing a detailed description of Internet information use in 7 consecutive cohorts of seriously ill patients from a specific set of hospitals that allows for the analysis of change over time. This is also a first effort to consider whether Internet use is linked to differences between the health care organizations.

Women diagnosed with breast cancer have wide-ranging information needs. In a study of Internet savvy younger women (younger than 45 years) diagnosed with breast cancer, results suggested that these patients searched for information to help them make good treatment decisions, to learn about their future care and prospects, and to pursue social support [61]. These goals are congruent with principles of patient empowerment and involvement in health care decision making.

This study also has implications for practice that has not yet fully harnessed the healing and empowerment potential of technology for the benefit of persons living with life-threatening illnesses. Access to the information on the Internet has been shown to enhance health-promoting behaviors [62,63].

Acknowledgments

We thank the patients who participated in this study and the breast centers which supported this survey. We can confirm that all patient identifiers have been removed or disguised so that the patients described are not identifiable. The study was approved by the Medical Ethics Committee of the University Hospital of Cologne. The patient survey was requested and initiated by the North Rhine-Westfalian Ministry of Work, Health and Social Affairs (MAGS NRW). Costs were borne by the participating hospitals as part of the breast center certification and benchmarking process. The hospitals asked the patients to consent to participation in the study and indicated disease characteristics as reported. Neither the ministry nor the hospitals were involved in the analysis or interpretation of data, in the writing of the report, or in the decision to submit the paper for publication. Parts of the data analysis and the writing were done during the first author's stay at the University of Michigan that was financed by the German Research Foundation (Grant number: Ko 3856/2). Many thanks to Markus Alich, the care research laboratory manager at the Institute for Medical Sociology, Health Services Research and Rehabilitation Science. Parts of the results were presented at the 2013 International Conference on Communication in Healthcare in Montréal, Canada.

Conflicts of Interest

None declared.

References

1. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
2. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
3. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
4. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004 Sep 30;6(3):e36 [FREE Full text] [doi: 10.2196/jmir.6.3.e36] [Medline: 15471762]
5. Eisemann N, Waldmann A, Katalinic A. Epidemiology of Breast Cancer - Current Figures and Trends. Geburtshilfe Frauenheilkd 2013 Feb;73(2):130-135 [FREE Full text] [doi: 10.1055/s-0032-1328075] [Medline: 24771909]
6. Seckin G. Patients as information managers: the internet for successful self-health care and illness management. Open Longev Sci 2010;4:36-42. [doi: 10.2174/1876326X01004010036]
7. Lee CJ, Gray SW, Lewis N. Internet use leads cancer patients to be active health care consumers. Patient Educ Couns 2010 Dec;81 Suppl:S63-S69. [doi: 10.1016/j.pec.2010.09.004] [Medline: 20889279]
8. Seckin G. Cyber patients surfing the medical web: Computer-mediated medical knowledge and perceived benefits. Comp Hum Behav 2010;26(6):1694-1700. [doi: 10.1016/j.chb.2010.06.018]
9. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
10. Eysenbach G. The impact of the Internet on cancer outcomes. CA Cancer J Clin 2003;53(6):356-371. [Medline: 15224975]
11. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. Int J Med Inform 2006 Jan;75(1):29-41. [doi: 10.1016/j.ijmedinf.2005.07.029] [Medline: 16129659]
12. Wald HS, Dube CE, Anthony DC. Untangling the Web--the impact of Internet use on health care and the physician-patient relationship. Patient Educ Couns 2007 Nov;68(3):218-224. [doi: 10.1016/j.pec.2007.05.016] [Medline: 17920226]
13. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
14. Helft PR, Hlubocky F, Daugherty CK. American oncologists' views of internet use by cancer patients: a mail survey of American Society of Clinical Oncology members. J Clin Oncol 2003 Mar 1;21(5):942-947. [Medline: 12610198]
15. Lewis T. Seeking health information on the internet: lifestyle choice or bad attack of cyberchondria? Media, Culture & Society 2006 Jul 01;28(4):521-539 [FREE Full text] [doi: 10.1177/0163443706065027]
16. Chung JE. Patient-provider discussion of online health information: results from the 2007 Health Information National Trends Survey (HINTS). J Health Commun 2013;18(6):627-648. [doi: 10.1080/10810730.2012.743628] [Medline: 23590202]
17. Bylund CL, Gueguen JA, Sabee CM, Imes RS, Li Y, Sanford AA. Provider-patient dialogue about Internet health information: an exploration of strategies to improve the provider-patient relationship. Patient Educ Couns 2007 Jun;66(3):346-352. [doi: 10.1016/j.pec.2007.01.009] [Medline: 17329062]
18. McMullan M. Patients using the Internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]

19. Lawrentschuk N, Sasges D, Tasevski R, Abouassaly R, Scott AM, Davis ID. Oncology health information quality on the Internet: a multilingual evaluation. Ann Surg Oncol 2012 Mar;19(3):706-713. [doi: 10.1245/s10434-011-2137-x] [Medline: 22146882]
20. National Institutes of Health. Health information URL: http://health.nih.gov/ [accessed 2014-07-30] [WebCite Cache ID 6RSJsVTY4]
21. Agency for Healthcare Research and Quality. Topics A-Z URL: http://www.ahrq.gov/health-care-information/ [accessed 2014-07-30] [WebCite Cache ID 6RSK85vYI]
22. National Cancer Institute. Types of cancer URL: http://www.cancer.gov/ [accessed 2014-07-30] [WebCite Cache ID 6RSKC2HO1]
23. gesundheitsinformation.de. URL: http://www.gesundheitsinformation.de/ [accessed 2014-07-30] [WebCite Cache ID 6RSKX7lrk]
24. Krebsgesellschaft. URL: http://krebsgesellschaft.de/ [WebCite Cache ID 6RSWlbIa1]
25. Littlechild SA, Barr L. Using the Internet for information about breast cancer: a questionnaire-based study. Patient Educ Couns 2013 Sep;92(3):413-417. [doi: 10.1016/j.pec.2013.06.018] [Medline: 23891419]
26. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Use of the Internet by women with breast cancer. J Med Internet Res 2002;4(2):E9 [FREE Full text] [doi: 10.2196/jmir.4.2.e9] [Medline: 12554556]
27. Angelucci E, Orlando A, Ardizzone S, Guidi L, Sorrentino D, Fries W, et al. Internet use among inflammatory bowel disease patients: an Italian multicenter survey. Eur J Gastroenterol Hepatol 2009 Sep;21(9):1036-1041. [doi: 10.1097/MEG.0b013e328321b112] [Medline: 19543105]
28. van Uden-Kraan CF, Drossaert CH, Taal E, Smit WM, Moens HJ, Siesling S, et al. Health-related Internet use by patients with somatic diseases: frequency of use and characteristics of users. Inform Health Soc Care 2009 Jan;34(1):18-29. [doi: 10.1080/17538150902773272] [Medline: 19306196]
29. van de Poll-Franse LV, van Eenbergen MC. Internet use by cancer survivors: current use and future wishes. Support Care Cancer 2008 Oct;16(10):1189-1195. [doi: 10.1007/s00520-008-0419-z] [Medline: 18293014]
30. Satterlund MJ, McCaul KD, Sandgren AK. Information gathering over time by breast cancer patients. J Med Internet Res 2003;5(3):e15 [FREE Full text] [doi: 10.2196/jmir.5.3.e15] [Medline: 14517106]
31. Bianco A, Zucco R, Nobile CG, Pileggi C, Pavia M. Parents seeking health-related information on the Internet: cross-sectional study. J Med Internet Res 2013;15(9):e204 [FREE Full text] [doi: 10.2196/jmir.2752] [Medline: 24047937]
32. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
33. Weaver JB, Mays D, Lindner G, Eroglu D, Fridinger F, Bernhardt JM. Profiling characteristics of internet medical information users. J Am Med Inform Assoc 2009;16(5):714-722 [FREE Full text] [doi: 10.1197/jamia.M3150] [Medline: 19567794]
34. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243. [doi: 10.1177/1460458211414843] [Medline: 21937464]
35. Hsu J, Huang J, Kinsman J, Fireman B, Miller R, Selby J, et al. Use of e-Health services between 1999 and 2002: a growing digital divide. J Am Med Inform Assoc 2005;12(2):164-171 [FREE Full text] [doi: 10.1197/jamia.M1672] [Medline: 15561786]
36. Kowalski C, Würstlein R, Steffen P. : Vier Jahre Patientinnenbefragung im Rahmen der (Re-)Zertifizierung der Brustzentren in Nordrhein-Westfalen. Senologie 2010;7:24-28. [doi: 10.1055/s-0030-1247275]
37. Dillman DA. Mail and Telephone Surveys: The Total Design Method. New York: Wiley; 1978.
38. Sobin LH, Wittekind C. UICC TNM Classification of Malignant Tumors. 6th edition. New York: John Wiley & Sons; 2002.
39. Hox JJ. Multilevel Analysis: Techniques and Applications. Mahwah, NJ: Lawrence Erlbaum Associates; 2002.
40. Hesse BW, Arora NK, Burke Beckjord E, Finney Rutten LJ. Information support for cancer survivors. Cancer 2008 Jun 1;112(11 Suppl):2529-2540 [FREE Full text] [doi: 10.1002/cncr.23445] [Medline: 18428201]
41. Houston TK, Allison JJ. Users of Internet health information: differences by health status. J Med Internet Res 2002;4(2):E7 [FREE Full text] [doi: 10.2196/jmir.4.2.e7] [Medline: 12554554]
42. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? Health Serv Res 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00510.x] [Medline: 16704514]
43. Wietfeldt ED, Hassan I, Johnson A, Veeder C, Boehler M, Dunnington GL, et al. Health-related information gathering practices in breast and colorectal specialty clinics: the end of the electronic divide? Am J Surg 2010 Mar;199(3):364-7; discussion 367. [doi: 10.1016/j.amjsurg.2009.09.018] [Medline: 20226911]
44. McCully SN, Don BP, Updegraff JA. Using the Internet to help with diet, weight, and physical activity: results from the Health Information National Trends Survey (HINTS). J Med Internet Res 2013;15(8):e148 [FREE Full text] [doi: 10.2196/jmir.2612] [Medline: 23906945]
45. Choi NG, Dinitto DM. Internet use among older adults: association with health needs, psychological capital, and social capital. J Med Internet Res 2013;15(5):e97 [FREE Full text] [doi: 10.2196/jmir.2333] [Medline: 23681083]

46. Hale TM, Cotton SR, Drentea P, Goldner M. Rural-Urban Differences in General and Health-Related Internet Use. Am Behav Sci 2010;53(9):1304-1325. [doi: 10.1177/0002764210361685]
47. Clayman ML, Manganello JA, Viswanath K, Hesse BW, Arora NK. Providing health messages to Hispanics/Latinos: understanding the importance of language, trust in health information sources, and media use. J Health Commun 2010;15 Suppl 3:252-263 [FREE Full text] [doi: 10.1080/10810730.2010.522697] [Medline: 21154097]
48. Mesch G, Mano R, Tsamir J. Minority status and health information search: a test of the social diversification hypothesis. Soc Sci Med 2012 Sep;75(5):854-858. [doi: 10.1016/j.socscimed.2012.03.024] [Medline: 22633160]
49. Sun GH. The digital divide in Internet-based patient education materials. Otolaryngol Head Neck Surg 2012 Nov;147(5):855-857. [doi: 10.1177/0194599812456153] [Medline: 22850179]
50. Eloy JA, Li S, Kasabwala K, Agarwal N, Hansberry DR, Baredes S, et al. Readability assessment of patient education materials on major otolaryngology association websites. Otolaryngol Head Neck Surg 2012 Nov;147(5):848-854. [doi: 10.1177/0194599812456152] [Medline: 22864405]
51. Percheski C, Hargittai E. Health information-seeking in the digital age. J Am Coll Health 2011;59(5):379-386. [doi: 10.1080/07448481.2010.513406] [Medline: 21500056]
52. Iredale R, Brain K, Williams B, France E, Gray J. The experiences of men with breast cancer in the United Kingdom. Eur J Cancer 2006 Feb;42(3):334-341. [doi: 10.1016/j.ejca.2005.09.027] [Medline: 16377181]
53. Williams BG, Iredale R, Brain K, France E, Barrett-Lee P, Gray J. Experiences of men with breast cancer: an exploratory focus group study. Br J Cancer 2003 Nov 17;89(10):1834-1836 [FREE Full text] [doi: 10.1038/sj.bjc.6601305] [Medline: 14612888]
54. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
55. Zulman DM, Kirch M, Zheng K, An LC. Trust in the internet as a health resource among older adults: analysis of data from a nationally representative survey. J Med Internet Res 2011 Feb;13(1):e19 [FREE Full text] [doi: 10.2196/jmir.1552] [Medline: 21324832]
56. Voelker R. Seniors seeking health information need help crossing "digital divide". JAMA 2005 Mar 16;293(11):1310-1312. [doi: 10.1001/jama.293.11.1310] [Medline: 15769951]
57. 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]
58. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
59. Xie B, Wang M, Feldman R, Zhou L. Internet use frequency and patient-centered care: measuring patient preferences for participation using the health information wants questionnaire. J Med Internet Res 2013;15(7):e132 [FREE Full text] [doi: 10.2196/jmir.2615] [Medline: 23816979]
60. Flynn KE, Smith MA, Freese J. When do older adults turn to the internet for health information? Findings from the Wisconsin Longitudinal Study. J Gen Intern Med 2006 Dec;21(12):1295-1301 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00622.x] [Medline: 16995892]
61. Balka E, Krueger G, Holmes, BJ, Stephen, JE. Situating internet use: Information-seeking among young women with breast cancer. J Comput-Mediat Commun 2010;15(3):389-411. [doi: 10.1111/j.1083-6101.2010.01506.x]
62. 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]
63. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. Patient Educ Couns 2010 Mar;78(3):329-336. [doi: 10.1016/j.pec.2010.01.013] [Medline: 20202779]

Abbreviations

ICC: intraclass correlation coefficient
ICD: International Classification of Diseases
UICC: Union for International Cancer Control

Edited by G Eysenbach; submitted 31.01.14; peer-reviewed by E Neter, C Zhan; comments to author 28.04.14; accepted 14.07.14; published 26.08.14.

Please cite as:
Kowalski C, Kahana E, Kuhr K, Ansmann L, Pfaff H
Changes Over Time in the Utilization of Disease-Related Internet Information in Newly Diagnosed Breast Cancer Patients 2007 to 2013
J Med Internet Res 2014;16(8):e195
URL: http://www.jmir.org/2014/8/e195/
doi:10.2196/jmir.3289
PMID:25158744

Original Paper

Motives for Participating in a Web-Based Nutrition Cohort According to Sociodemographic, Lifestyle, and Health Characteristics: The NutriNet-Santé Cohort Study

Caroline Méjean[1], PhD; Fabien Szabo de Edelenyi[1], PhD; Mathilde Touvier[1], PhD; Emmanuelle Kesse-Guyot[1], PhD; Chantal Julia[1], PhD; Valentina A Andreeva[1], PhD; Serge Hercberg[1], PhD

Université Paris 13, Sorbonne Paris Cité, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre d'Epidémiologie et Biostatistiques Paris Nord, Inserm (U1153), Inra (U1125), Cnam, Université Paris 5, Université Paris 7, Bobigny, France

Corresponding Author:
Caroline Méjean, PhD
Université Paris 13, Sorbonne Paris Cité, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre d'Epidémiologie et Biostatistiques Paris Nord, Inserm (U1153), Inra (U1125), Cnam, Université Paris 5, Université Paris 7
74 rue marcel cachin
Bobigny, F-93017
France
Phone: 33 148388958
Fax: 33 148388931
Email: c.mejean@uren.smbh.univ-paris13.fr

Abstract

Background: In traditional epidemiological studies, participants are likely motivated by perceived benefits, feelings of accomplishment, and belonging. No study has explored motives for participation in a Web-based cohort and the associated participant characteristics, although such information is useful for enhancing recruitment and improving cohort retention.

Objective: We aimed to evaluate the relationships between motives for participation and sociodemographic, health, and lifestyle characteristics of participants in the NutriNet-Santé Web-based cohort, designed to identify nutritional risk or protective factors for chronic diseases.

Methods: The motives for participation were assessed using a specifically developed questionnaire administered approximately 2 years after baseline. A total of 6352 completed the motives questionnaire (43.34%, 6352/15,000 randomly invited cohort participants). We studied the associations between motives (dependent variables) and individual characteristics with multivariate multinomial logistic regression models providing odds ratios and 95% confidence intervals.

Results: In total, 46.45% (2951/6352) of participants reported that they would not have enrolled if the study had not been conducted on the Internet, whereas 28.75% (1826/6352) were not sure. Men (OR 1.21, 95% CI 1.04-1.42), individuals aged 26-35 years (OR 1.51, 95% CI 1.20-1.91), and obese participants (OR 1.30, 95% CI 1.02-1.65) were more inclined to be motivated by the Internet aspect. Compared with younger adults and managerial staff, individuals >55 years (OR 0.60, 95% CI 0.48-0.45) and employees/manual workers were less likely motivated by the Internet aspect (OR 0.77, 95% CI 0.63-0.92). Regarding reasons for participation, 61.37% (3898/6352) reported participating to help advance public health research on chronic disease prevention; 22.24% (1413/6352) to help advance nutrition-focused research; 6.89% (438/6352) in response to the call from the media, after being encouraged by a close friend/associate, or a medical provider. Individuals >45 years (vs younger participants) were more likely (OR 1.62, 95% CI 1.07-2.47), whereas overweight and obese participants (vs nonobese participants) were less likely to participate in the study for reasons related to helping public health research on chronic disease prevention (OR 0.72, 95% CI 0.58-0.89; OR 0.62, 95% CI 0.46-0.84; respectively). Exclusive public funding of the study was important for 67.02% (4257/6352) of the participants. Men (OR 1.37, 95% CI 1.17-1.61) and persons >55 years (OR 1.97, 95% CI 1.57-2.47) were more likely to consider the exclusive public funding as very important.

Conclusions: The use of the Internet, the willingness to help advance public health research, and the study being publicly funded were key motives for participating in the Web-based NutriNet-Santé cohort. These motives differed by sociodemographic profile and obesity, yet were not associated with lifestyle or health status. These findings can help improve the retention strategies in Web-based cohorts, particularly during decisive study periods when principal exposure information is collected.

(J Med Internet Res 2014;16(8):e189) doi:10.2196/jmir.3161

KEYWORDS

Internet; cohort study; population characteristics; motivation; participation

Introduction

The successful implementation of very large population-based cohort studies involving collection of comprehensive, high-quality dietary, lifestyle, and health data is both a priority and a challenge in nutritional epidemiology [1,2]. Such observational studies face very high logistic costs and require substantial resources. The rapid and widespread increase in access to Internet has made this tool a viable and logical base for cohort studies because it presents advantages across all research stages [3,4]. In most industrialized countries, Internet access is greater than 50% and is still increasing [5]. In Europe, Internet users are becoming more representative of the general population, including older adults (42% of individuals older than 55 years are regular users) and people of low socioeconomic status (73% of individuals belonging to low socioprofessional categories are regular users) [6]. In France, in November 2009, 34.7 million French citizens (approximately 65 % of the population older than 11 years) were connected to the Internet during the previous month [7].

Yet Web-based prospective cohort studies are still in their infancy [8]. Whereas issues related to participation are crucial in epidemiological studies [9], participation in both Web-based cohorts and repeated-measures cross-sectional studies, and associated sociodemographic profiles have been rarely investigated [8,10-22]. A few studies have compared the sociodemographic and lifestyle characteristics of participants who used Web-based questionnaires with paper-and-pencil respondents [8,14]. The Black Women’s Health Study showed that Web-based questionnaire users were younger and had higher socioeconomic status than paper-and-pencil users, but no difference was observed in terms of lifestyle or health status [14]. These authors underlined the fact that low socioeconomic status and older age, which are markers of Web access, remain barriers to the use of Web-based questionnaires. In turn, the Millennium cohort study with military personnel has highlighted that Web responders were more likely to be male, younger, highly educated, obese, and smokers than were paper-and-pencil responders [8]. The results regarding sex, age, and educational level seem to be due to greater technological competence in these groups and greater access to the Internet, whereas findings on weight status and smoking may reflect subtle occupational differences, such as being employed in a sedentary work environment or an unhealthier lifestyle outside of work. Also, the ATTEMPT cohort has shown that participants had similar sociodemographic and lifestyle profiles to those found in the general population [16], whereas NINFEA cohort participation, regarding Italian pregnant women, was associated with being older, having a higher educational level, and being native Italian compared to the general population, probably due to population-specific occurrence [13]. In Web-based intervention studies with repeated measurements, women, older participants, nonsmokers, heavy alcohol consumers, and overweight participants were more likely to remain enrolled in Internet-delivered behavior change programs [17-22].

Voluntariness refers to the voluntary motivational nature of a person’s participation from the initial decision to participate through the course of the study, and is influenced by external and internal factors [23]. In traditional epidemiological studies, participants are motivated by the benefits they perceive, particularly the information and care received during the medical examinations, the positive feelings about oneself or an enhanced self-image generated by the act of participation, and the sense of loyalty and belonging associated with being part of the study [24-28]. To our knowledge, no study has explored motives for participation in a Web-based cohort and the associated participant characteristics, particularly during key periods of collection of exposure data. Such information would be useful in enhancing the recruitment of diverse population samples and in improving cohort retention.

The NutriNet-Santé study was launched in May 2009 in France to investigate multiple facets of the relationship between nutrition and health along with determinants of dietary behavior [29]. This is a large, Web-based, prospective, nutritional epidemiology cohort. To date, 156,734 volunteers aged >18 years have been included in the cohort. Participants are followed via a website specifically created for that purpose. The objective of the present study was to assess motives for participation in the NutriNet-Santé cohort after 2 years of participation, such as reasons for participation, the influence of the Internet support in participation, and the importance of public funding. We also explored sociodemographic, lifestyle, and heath characteristics associated with those motives.

Methods

Population

Participants were part of the NutriNet-Santé Study, a large Web-based prospective observational cohort. It is implemented in a general population targeting Internet-using adult volunteers aged 18 years and older. The design, methods, and rationale have been described elsewhere [29]. Eligible participants were recruited by different means. Initially, a vast multimedia campaign (television, radio, national and regional newspapers, posters, and Internet) called for volunteers and provided details about the study’s website [12]. The multimedia campaigns are repeated every 6 months. Recruitment information was maintained on a large number of websites (national institutions, city councils, private firms, Web organizations) and is regularly updated via professional channels (eg, doctors, pharmacists, dentists, business partners, municipalities). The key message delivered in the call for volunteers was:

> *Numerous scientific studies have highlighted the role of nutrition as a protective factor or a risk of many common diseases in France, as in all industrialized countries, such as cancer, cardiovascular diseases,*

> *obesity, type 2 diabetes, dyslipidemia, and hypertension. Nutrition is not the only determinant of these health problems. Indeed, genetic, biological, and environmental factors are involved in the onset of these diseases. To highlight the specific role of nutritional factors in health, the development of cohort studies with very large populations (group of participants followed for several years) is essential as they permit to accurately measure food intake, but also take into account other determinants, such as physical activity, weight, smoking, and family history of disease. The purpose of our study is to identify nutritional risk factors or protective factors for these diseases, which is an essential step in establishing dietary recommendations to prevent the risk of disease and improve the health of the current and future generations. This is the ambitious goal of the NutriNet-Santé study and that is why researchers need you.*

During each multimedia recruitment campaign and during the enrollment process, participants are informed that follow-up over at least 10 years is planned.

Previous findings showed that most of the participants enrolled after hearing about the study on television because this medium entails the widest reach [12]. In particular, television announcements permitted the recruitment of members of population subgroups that are not typically well represented in population-based epidemiological research, those belonging to lower socioeconomic strata. The radio, newspapers, Internet, and advice from acquaintances also proved to be substantial means of disseminating information about this epidemiological study to encourage participation.

To be included, participants have to fill in on the website an initial set of questionnaires assessing dietary intake, physical activity, anthropometrics, lifestyle, and socioeconomic conditions along with health status. Participants were informed by email that, after inclusion, they would be asked to complete the same questionnaires each year as part of their follow-up. In addition, they are invited to fill in a complementary questionnaire each month. Aspects related to convenience of participation (ie, ≤20 min each month) and confidentiality were also emphasized. In addition, a system of boosting motivation and retention was implemented. In order to forge a sense of community that helps advance research, participants receive a NutriNet-Santé membership card at inclusion and a certificate on completion of each follow-up year/wave. They also receive monthly email with scientific information regarding health and nutrition, and invitations to press conferences about the study results. For purposes of retention, free screening tests for cholesterol, triglycerides, and diabetes are offered to participants (the results are sent back with a special notice in case of abnormal test results).

All baseline questionnaires were first pilot-tested and compared with traditional administration methods (paper-and-pencil versions or interviews by a dietitian) [30-32]. Health events are monitored through questionnaires about hospitalizations and medication use as well as via a linkage with the national vital statistics database. In April 2011, 15,000 randomly selected participants among 86,652 individuals included at the time of the study were invited to complete a Web-based questionnaire regarding their motives for participation in the study.

This study was conducted according to guidelines laid down in the Declaration of Helsinki, and all procedures were approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm no: 0000388FWA00005831) and the Commission Nationale Informatique et Libertés (CNIL no: 908450 and no: 909216). Written electronic informed consent to participate in the study was obtained from all participants.

Data Collection

Motives for Participation

Participants were asked, “Would you have participated in the NutriNet-Santé study if it were not Internet-based?” (response options: yes, no, I don’t know). We also asked the participants, “What was your main reason for participating in the NutriNet-Santé study?” The response options for the different motives were classified into 2 general categories: (1) intrinsic motives for participation, including, to help advance public health research on chronic disease prevention, to help advance nutrition research, to receive regular scientific information about health and nutrition, out of curiosity, to belong to a group, or other motives and (2) extrinsic motives, including in response to the call from the media, from a close friend/associate, or from a medical provider. Finally, we asked participants, “Is the fact that the study is exclusively funded by public sources important for your participation?” (response options: very important, important, not very important, not important).

Assessment of Individual Characteristics

At baseline, sociodemographic, lifestyle, and health characteristics were self-reported. Participants indicated their alcohol consumption frequency and quantity over the previous 7 days. Alcohol intake was calculated by multiplying the alcohol content (ie, percentage) of each beverage (wine, beer, spirits, and cider) by the standard ethanol weight content. Body mass index (BMI) was assessed using self-reported height and weight. Status regarding type 2 diabetes, hypertension, and hypercholesterolemia was provided by participants by answering the following question: “Have you been or are you currently being treated for type 2 diabetes / hypertension / hypercholesterolemia?” If the participant answered yes, he/she completed the information by self-reporting the year of diagnosis and current use of medication.

Statistical Analysis

The present analyses focused on data from a random sample of participants in the NutriNet-Santé cohort who had completed the questionnaire assessing their participation motives and who had no missing sociodemographic, lifestyle, anthropometric, or health status data. These characteristics were compared between participants included in our analysis and those who had stopped participating within 6 months after their enrollment (calculated from the date of the last connection on the website), using a chi-square goodness-of-fit test. The possible reasons for participation were grouped into the following 4 categories: (1) to help advance public health research on chronic disease prevention, (2) to help advance nutrition research, (3) in response to the call from the media, from a close friend/associate or from a medical provider, and (4) other motives (ie, to receive regular scientific information about health and nutrition, out of curiosity, to belong to a group, and other). Perceptions/attitudes toward the public funding of the study were categorized into 3 groups: very important, important, and not important.

According to French recommendations [33], male drinkers were categorized as abstainers and irregular consumers (<once a week), moderate (0-30 g alcohol/day), or heavy drinkers (>30 g alcohol/day), and females as nondrinkers (0 g alcohol/day), moderate (0-20 g alcohol/day), or heavy drinkers (>20 g alcohol/day). Normal weight, overweight, and obesity were defined according to the World Health Organization classification for BMI, as BMI <25 kg/m^2, 25≤ BMI <30 kg/m^2, and BMI ≥30 kg/m^2, respectively [34]. Gender, age (≤25, 26-35, 36-45, 46-55, and >55 years), marital status (married or living with a partner, single/divorced/widowed), having at least 1 child at home, education (elementary school, secondary, college graduate, and advanced degree), occupational category (managerial staff, farmers/self-employed, intermediate profession, employees/manual workers, and never-employed/homemaker), type of area of residence (rural, semiurban <20,000 inhabitants, urban 20,000-100,000 inhabitants, urban ≥100,000 inhabitants, Paris metropolitan area), smoking status (never smoker, former smoker, current smoker), alcohol consumption, BMI, self-reported type 2 diabetes, hypertension, and hypercholesterolemia are presented in a frequency/percent format for the whole sample. Multivariate associations between the individual characteristics and the motives for participation (participation motive related to the advantages of the Internet, reasons for participation, and attitudes about public funding) were assessed using multinomial logistic regression. Reference category used in the analysis of the associations between participation motive related to the advantages of the Internet and individual characteristics was yes. For the analysis regarding reasons for participation, the reference category was other motives, and for those concerning attitudes about public funding the reference was not important. In each multivariate model, the explanatory variables were adjusted for each other. Odds ratios (OR) and 95% confidence intervals (CI) are reported. Significance tests were 2-sided with a *P* value set at <.05. A more conservative *P* value of.01 was also used for estimating the robustness of the results. Statistical analyses were performed using SAS software version 9.3 (SAS Institute Inc, Cary, NC, USA).

Results

A total of 6556 of 15,000 persons completed the motives questionnaire (ie, 43.71% of the randomly invited cohort participants). We excluded 61 individuals with missing data regarding the socioeconomic characteristics, 135 participants with missing data regarding weight or height, and 11 participants with missing data regarding alcohol consumption; therefore, data from 6352 participants was available for analysis. At the time of the administration of the questionnaire about motives, the mean duration of participation in the cohort for the participants included in this analysis was 20 months (SD 4.00) and the median was 23 months (range 1-24). Characteristics of the sample are presented in Table 1.

Compared with nonrespondents (among the 15,000 contacted participants), the percentages of individuals older than 55 years and of managerial staff were higher among participants included in this analysis, whereas the percentage of individuals with at least 1 child at home was lower (data not shown). Compared with participants who stopped participating within 6 months after their inclusion in the cohort (mean duration of participation: 3 weeks after inclusion, SD 1 week), the percentages of individuals older than 45 years, of married persons, managerial staff, persons with high educational level, individuals who reported hypertension, and those who reported hypercholesterolemia were higher among participants included in this analysis, whereas the percentages of individuals with at least 1 child at home, manual workers/employees, infrequent alcohol consumers, smokers, and obese individuals were lower (Table 1).

Among participants, 46.45% (2951/6352) reported that they would not have enrolled had the study not been conducted on the Internet, whereas 28.75% (1826/6352) were not sure (Table 2).

Table 1. Characteristics of the sample.[a]

Individual characteristics	Present sample, n (%) n=6352	Drop-out,[b] n (%) n=9982
Gender		
Female	4821 (75.90)	7584 (75.98)
Male	1531 (24.10)	2398 (24.02)
Age (years)		
18-25	480 (7.56)	1482 (14.85)
26-35	1133 (17.84)	2630 (26.35)
36-45	1211 (19.06)	2276 (22.80)
46-55	1344 (21.16)	1843 (18.46)
>55	2184 (34.38)	1750 (17.54
Marital status		
Married or living with a partner	4680 (73.68)	6739 (67.51)
Single, divorced, widowed	1672 (26.32)	3243 (32.49)
Have at least 1 child at home		
Yes	1976 (31.11)	3946 (39.53)
No	4376 (68.89)	6036 (60.47)
Educational level		
Advanced/graduate degree	2031 (31.98)	2414 (24.18)
College graduate	1868 (29.41)	2856 (28.61)
Secondary	2233 (35.15)	4261 (42.69)
Elementary school	220 (3.46)	451 (4.52)
Occupational category		
Managerial staff	2215 (34.87)	2437 (24.41)
Self-employed, farmer	198 (3.12)	424 (4.25)
Intermediate profession	1673 (26.34)	2101 (21.05)
Employee, manual worker	1959 (30.84)	4389 (43.97)
Never-employed/homemaker	307 (4.83)	631 (6.32)
Area of residence		
Rural	1393 (21.98)	2014 (20.18)
Semiurban, population <20,000	997 (15.71)	1445 (14.48)
Urban, population between 20,000-100,000	784 (12.39)	1243 (12.45)
Urban, population ≥100,000	2118 (33.23)	3408 (34.14)
Urban, Paris	1060 (16.69)	1872 (18.75)
Alcohol consumption		
Abstainers and infrequent consumers (<once a week)	1635 (25.74)	2920 (29.25)
Moderate consumption (≤20 g/day for women and ≤30 g/day for men)	4192 (65.99)	6308 (63.19)
Heavy consumption (>20 g/day for women and >30 g/day for men)	525 (8.27)	754 (7.55)
Smoking status		
Never smoker	3195 (50.30)	4374 (43.82)
Former smoker	2200 (34.63)	2858 (28.63)

Individual characteristics	Present sample, n (%) n=6352	Drop-out,[b] n (%) n=9982
Current smoker	957 (15.07)	2750 (27.55)
Weight status (BMI)		
Normal (<25 kg/m^2)	4410 (69.43)	6461 (64.73)
Overweight (≥25 kg/m^2-30 kg/m^2>)	1382 (21.76)	2262 (22.66)
Obese (≥30 kg/m^2)	560 (8.82)	1259 (12.61)
Self-reported type 2 diabetes (yes)	157 (2.47)	206 (2.06)
Self-reported hypertension (yes)	876 (13.79)	982 (9.84)
Self-reported hypercholesterolemia (yes)	755 (11.89)	678 (6.79)

[a] All *P* values were <.05.

[b] Individuals who stopped participating within 6 months after their inclusion in the cohort.

Table 2. Motives for participation in the NutriNet-Santé cohort study (N=6352).

Motives for participation	n (%)
What was your main reason to participate in the study?	
To help advance nutrition research	1413 (22.24)
To help advance public health research on chronic disease prevention	3898 (61.37)
In response to the call for volunteers (from media, a friend/associate or a medical provider)	438 (6.89)
Other[a]	603 (9.50)
Would you have participated in the study if it were not Internet-based?	
Yes	1575 (24.80)
No	2951 (46.45)
Don’t know	1826 (28.75)
Is the fact that the study is exclusively funded by public sources important for your participation?	
Very important	2185 (34.40)
Important	2072 (32.62)
Not important	2095 (32.98)

[a] Other category includes participation to receive regular scientific information about health and nutrition, out of curiosity, to belong to a group, and other.

Compared to women, men were more inclined to be motivated by the Internet aspect (Table 3). Individuals aged between 26 and 35 years, those without children at home and obese persons also displayed increased motivation associated with the Internet aspect. Individuals older than 55 years, employees/manual workers, and those in intermediate professions were less likely to be motivated by the Internet aspect compared to younger adults and managerial staff (Table 3). Results regarding occupational categories and having at least 1 child at home did not remain significant when the more conservative *P* value of.01 was used.

Table 3. Sociodemographic, lifestyle, and health characteristics associated with motives for participation in the study had it not been Internet-based (multivariate analysis, N=6352)

Individual characteristics	No, I would not have enrolled[a]		I don't know[a]	
	OR	95% CI	OR	95% CI
Gender				
Female	1.00		1.00	
Male	1.22	1.04-1.43	0.96	0.80-1.14
Age (years)				
18-25	1.21	0.86-1.70	1.19	0.82-1.73
26-35	1.51	1.20-1.91	1.39	1.08-1.79
36-45	1.00		1.00	
46-55	0.90	0.74-1.11	0.86	0.68-1.07
>55	0.61	0.49- 0.76	0.76	0.59-0.96
Marital status				
Married or living with a partner	1.00		1.00	
Single, divorced, widowed	0.97	0.83-1.13	1.00	0.85-1.18
Have at least 1 child at home				
Yes	1.00		1.00	
No	1.23	1.04-1.46	1.22	1.02-1.47
Educational level				
Advanced/graduate degree	1.00		1.00	
College graduate	1.10	0.92-1.31	0.98	0.81-1.19
Secondary	1.10	0.91-1.33	1.00	0.82-1.24
Elementary school	1.18	0.80-1.74	1.38	0.92-2.07
Occupational category				
Managerial staff	1.00		1.00	
Self-employed, farmer	1.09	0.74-1.60	1.14	0.75-1.75
Intermediate profession	0.80	0.67-0.96	1.04	0.85-1.26
Employee, manual worker	0.77	0.63-0.93	0.89	0.72-1.10
Never-employed/homemaker	1.03	0.69-1.55	1.21	0.78-1.87
Area of residence				
Rural	1.00		1.00	
Semiurban population <20,000	1.01	0.83-1.23	1.07	0.85-1.34
Urban, population between 20,000-100,000	0.81	0.65-1.01	0.85	0.67-1.08
Urban, population ≥100,000	0.91	0.77-1.08	0.98	0.81-1.18
Urban, Paris	1.00	0.81-1.23	1.03	0.82-1.29
Alcohol consumption				
Abstainers and infrequent consumers (<once a week)	1.00			
Moderate consumption (≤ 20 g/day for women and ≤30 g/day for men)	1.18	1.02-1.37	1.20	1.02-1.40
Heavy consumption (>20 g/day for women and >30 g/day for men)	1.19	0.92-1.53	0.98	0.74-1.31
Smoking status				
Never smoker	1.00		1.00	
Former smoker	0.97	0.84-1.12	1.01	0.86-1.18

Individual characteristics	No, I would not have enrolled[a]		I don't know[a]	
	OR	95% CI	OR	95% CI
Current smoker	0.94	0.78-1.13	1.01	0.82-1.24
Weight status (BMI)				
Normal (<25 kg/m^2)	1.00		1.00	
Overweight (≥25 kg/m^2-30 kg/m^2>)	1.07	0.91-1.25	1.09	0.92-1.30
Obese (≥30 kg/m^2)	1.32	1.04-1.65	1.44	1.11-1.86
Self-reported type 2 diabetes (yes)	0.90	0.61-1.34	0.70	0.44-1.10
Self-reported hypertension (yes)	1,06	0.87-1.28	0.92	0.74-1.14
Self-reported hypercholesterolemia (yes)	1.05	0.86-1.28	1.04	0.83-1.30

[a] The question was "Would you have participated in the study if it were not Internet-based?" Reference category for the outcome variable was "Yes, I would still participate even if the study was not Internet-based."

Regarding reasons for participation, 61.37% (3898/6352) reported participating to help advance public health research on chronic disease prevention; 22.24% (1413/6352) to help advance nutrition-focused research; 6.89% (438/6352) in response to a call from the media, a close friend/relative, or a medical professional; and 9.50% (603/6352) for other reasons (Table 2). Unlike younger participants, individuals older than 45 years were more likely to participate to help advance public health research on chronic disease prevention and to help advance nutrition research (Table 4). Overweight or obese persons were less inclined to participate for these reasons compared with individuals with normal weight. Single, divorced, or widowed individuals were less likely to participate for reasons of helping nutrition research or in response to the call from the media, a close friend/relative, or a medical professional than were individuals living with a partner. Finally, never-employed/homemakers were less likely to participate to help advance public health research on chronic disease prevention than were managerial staff.

Exclusive public funding for the study was important for two-thirds of the participants. Among them, half (2185/6352, 34.40%) considered it as very important (Table 2). Compared to women and to individuals aged between 36 and 45 years, men and persons older than 45 years were more likely to consider the exclusive public funding as very important or important, whereas younger participants were less likely to find it very important or important (Table 5). Compared to persons with the highest educational level, managerial staff and those with at least 1 child at home, individuals with less formal education, self-employed/farmers, employees/manual workers, or those without children at home were less likely to consider the exclusive public funding as very important or important. Results regarding having at least 1 child at home did not remain significant when the more conservative *P* value of .01 was used.

Table 4. Sociodemographic, lifestyle, and health characteristics associated with reasons for participation in the study (multivariate analysis, N=6352).

Individual characteristics	To help advance public research on chronic disease prevention[a]		To help advance nutrition research[a]		In response to the call (from the media, a friend/associate or a medical provider)[a]	
	OR	95% CI	OR	95% CI	OR	95% CI
Gender						
Female	1.00		1.00		1.00	
Male	1.02	0.82-1.27	0.86	0.67-1.09	0.96	0.70-1.32
Age (years)						
18-25	1.00		1.00		1.00	
26-35	1.08	0.72-1.61	1.43	0.91-2.25	1.05	0.59-1.89
36-45	1.31	0.85-2.01	1.55	0.96-2.51	1.00	0.53-1.87
46-55	1.63	1.07-2.48	1.74	1.09-2.79	1.48	0.81-2.69
> 55	1.62	1.07-2.46	1.43	0.90-2.29	1.33	0.73-2.41
Marital status						
Married or living with a partner	1.00		1.00		1.00	
Single, divorced, widowed	0.82	0.67-1.01	0.79	0.63-0.99	0.69	0.51-0.94
Have at least 1 child at home						
Yes	1.00		1.00		1.00	
No	1.02	0.82-1.29	0.9	0.70-1.15	1.08	0.78-1.49
Educational level						
Advanced/graduate degree	1.00		1.00		1.00	
College graduate	1.01	0.79-1.28	1.17	0.90-1.52	1.17	0.82-1.66
Secondary	1.13	0.86-1.47	0.97	0.72-1.31	1.21	0.82-1.78
Elementary school	1.39	0.80-2.40	0.66	0.34-1.29	1.26	0.59-2.72
Occupational category						
Managerial staff	1.00		1.00		1.00	
Self-employed, farmer	1.51	0.79-2.89	1.51	0.75-3.04	2.23	0.99-5.00
Intermediate profession	0.97	0.75-1.25	0.89	0.68-1.18	1.02	0.71-1.48
Employee, manual worker	0.85	0.65-1.10	0.75	0.56-1.00	1.14	0.78-1.67
Never-employed/homemaker	0.56	0.35-0.91	0.77	0.46-1.31	0.98	0.49-1.96
Area of residence						
Rural	1.00		1.00		1.00	
Semiurban, population <20,000	1.28	0.95-1.72	1.15	0.83-1.60	1.03	0.68-1.56
Urban, population between 20,000-100,000	0.99	0.73-1.34	1.03	0.73-1.44	1.01	0.66-1.56
Urban, population ≥100,000	0.96	0.76-1.21	0.95	0.73-1.23	0.91	0.65-1.28
Urban, Paris	1.28	0.95-1.72	1.07	0.77-1.48	1.12	0.74-1.70
Alcohol consumption						
Abstainers and infrequent consumers (<once a week)	1.00		1.00		1.00	
Moderate consumption (≤20 g/day for women and ≤30 g/day for men)	0.84	0.68-1.03	0.84	0.66-1.06	0.80	0.60-1.08
Heavy consumption (>20 g/day for women and >30 g/day for men)	0.85	0.59-1.22	0.84	0.56-1.26	0.80	0.48-1.34
Smoking status						

Individual characteristics	To help advance public research on chronic disease prevention[a]		To help advance nutrition research[a]		In response to the call (from the media, a friend/associate or a medical provider)[a]	
	OR	95% CI	OR	95% CI	OR	95% CI
Never smoker	1.00		1.00		1.00	
Former smoker	0.88	0.72-1.08	0.83	0.66-1.04	0.83	0.62-1.11
Current smoker	0.91	0.71-1.18	0.96	0.72-1.27	1.02	0.71-1.47
Weight status (BMI)						
Normal (<25 kg/m^2)	1.00		1.00		1.00	
Overweight (≥25 kg/m^2-30 kg/m^2>)	0.72	0.58-0.89	0.72	0.57-0.92	0.94	0.69-1.28
Obese (≥30 kg/m^2)	0.62	0.46-0.84	0.71	0.50-1.00	0.80	0.52-1.24
Self-reported type 2 diabetes (yes)	1.77	0.87-3.60	1.59	0.73-3.49	2.05	0.84-5.01
Self-reported hypertension (yes)	0.83	0.63-1.09	0.81	0.59-1.11	0.83	0.56-1.23
Self-reported hypercholesterolemia (yes)	1.30	0.95-1.77	1.23	0.87-1.74	1.08	0.70-1.68

[a] The question was “What was your main reason to participate in the study?” Reference category for the outcome variable was “other reasons” which includes participation to receive regular scientific information about health and nutrition, out of curiosity, to belong to a group, and other.

Table 5. Sociodemographic, lifestyle, and health characteristics associated with importance for exclusive public funding (multivariate analysis, N=6352).

Individual characteristics	Very important[a]		Important[a]	
	OR	95% CI	OR	95% CI
Gender				
Female	1.00		1.00	
Male	1.37	1.17-1.61	1.21	1.03-1.42
Age (years)				
18-25	0.40	0.28-0.58	0.57	0.41-0.78
26-35	0.63	0.51-0.79	0.79	0.64-0.97
36-45	1.00		1.00	
46-55	1.49	1.21-1.83	1.14	0.93-1.40
> 55	1.97	1.57-2.47	1.37	1.09-1.71
Marital status				
Married or living with a partner	1.00		1.00	
Single, divorced, widowed	0.96	0.82-1.12	1.09	0.94-1.26
Have at least 1 child at home				
Yes	1.00		1.00	
No	0.81	0.68-0.95	0.94	0.80-1.11
Educational level				
Advanced/graduate degree	1.00		1.00	
College graduate	0.80	0.67-0.96	0.92	0.77-1.10
Secondary	0.41	0.34-0.50	0.61	0.50-0.74
Elementary school	0.38	0.26-0.57	0.63	0.44-0.92
Occupational category				
Managerial staff	1.00		1.00	
Self-employed, farmer	0.63	0.43-0.93	0.87	0.61-1.26
Intermediate profession	0.87	0.73-1.04	0.90	0.75-1.08
Employee, manual worker	0.54	0.44-0.65	0.71	0.59-0.86
Never-employed/homemaker	0.76	0.50-1.15	1.16	0.81-1.66
Area of residence				
Rural	1.00		1.00	
Semiurban, population <20,000	1.02	0.83-1.26	1.06	0.86-1.30
Urban, population between 20,000-100,000	0.95	0.76-1.18	0.93	0.75-1.16
Urban, population ≥100,000	1.08	0.91-1.29	1.04	0.88-1.24
Urban, Paris	0.97	0.78-1.19	1.01	0.82-1.24
Alcohol consumption				
Abstainers and infrequent consumers <once a week)	1.00		1.00	
Moderate consumption ≤20 g/day for women and ≤30 g/day for men)	0.99	0.86-1.16	0.97	0.84-1.13
Heavy consumption >20 g/day for women and >30 g/day for men)	0.78	0.60-1.02	1.00	0.78-1.29
Smoking status				
Never smoker	1.00		1.00	
Former smoker	1.08	0.93-1.25	1.12	0.97-1.30
Current smoker	1.12	0.93-1.35	1.03	0.86-1.24

Individual characteristics	Very important[a]		Important[a]	
	OR	95% CI	OR	95% CI
Weight status (BMI)				
Normal <25 kg/m^2)	1.00		1.00	
Overweight ≥25 kg/m^2-30 kg/m^2>)	0.89	0.76-1.05	0.92	0.78-1.08
Obese ≥ 30 kg/m^2)	0.80	0.63-1.01	0.78	0.62-0.98
Self-reported type 2 diabetes (yes)	1.06	0.69-1.61	0.92	0.60-1.42
Self-reported hypertension (yes)	1.16	0.94-1.42	1.13	0.92-1.39
Self-reported hypercholesterolemia (yes)	1.13	0.91-1.40	1.18	0.95-1.46

[a] The question was "Is the fact that the study is exclusively funded by public sources important for your participation?" Reference category for the outcome variable was "not important."

Discussion

Principal Results

Profiles of participants in a Web-based epidemiological cohort have rarely been studied [8,10-16] and motives to participate in such cohorts have not yet been explored. The present study highlighted specific sociodemographic and health status (ie, weight status) characteristics of participants in a large Web cohort according to the perceived importance of the Internet, the reasons for participation, and the importance of public funding for research. Our results could guide the development of novel strategies for the retention of diverse population samples in Web-based cohorts, particularly during key periods of data collection.

Our findings revealed that almost half of the participants reported that Internet use was a decisive reason for participation. In fact, this element exerted a stronger influence among men, persons younger than 35 years, individuals of higher socioeconomic status, those without children at home, moderate alcohol consumers, and obese persons. Our results are concordant with previous studies that compared sociodemographic and lifestyle characteristics of participants who used Web-based questionnaires with those of participants who used paper-and-pencil instruments [8,14]. There is clear evidence that men, young people, and single persons are less likely to participate in epidemiological studies than are women, older, or married individuals [9,35-37]. Next, several studies have shown that individuals who presented risk behaviors, such as smoking, heavy alcohol use, or obesity, were often underrepresented among research participants [9,38,39]. Thus, our study suggests that the Internet allows for reaching a large population, but also provides access to hard-to-reach individuals given their social or behavioral status and for whom the Internet seems to be a more attractive and more convenient mean for participation compared to traditional means. Individuals belonging to low socioeconomic strata are often poorly represented when relying on traditional methods [40]. Our findings showed that the Internet appeared to be a less important motive for participation among individuals in low occupational categories, compared to managerial staff. Therefore, further exploration of measures that can be adopted in epidemiological Web-based studies to specifically increase opportunities for participation among low socioeconomic groups is necessary.

In addition, our study indicated that participation in an exclusively Web-based nutrition cohort was driven mainly by a desire to help advance research on chronic disease prevention or nutrition, especially among older participants, those with normal weight, and those who lived with a partner. Our results are also consistent with existing knowledge regarding motives for participation in volunteer-based cohorts on health and nutrition, which do not use the Internet [24,26-28]. These reasons for participation may reflect altruistic tendencies, but also a vested personal interest [27]. As explained by Williams et al [28], individuals may be more willing to participate if they believe that the potential benefits of their participation are large (eg, life versus death), highly likely to materialize, quickly attained, or likely to benefit themselves or someone important to them. Our results suggest that the desire to contribute to chronic disease risk prevention should be used as a key lever for participants' retention in Web-based cohorts, particularly during decisive periods of exposure data collection.

Two-thirds of our sample found the use of Internet for completion of the questionnaires to be a benefit, given its flexibility, whereas less than 1% found it to be complex. Also, one-quarter of participants felt more comfortable filling in the questionnaires on the website rather than face to face with an investigator. On the other hand, only 22% visited sections of the website of the study regarding news and progress of the study. Thus, in Web-based studies, the reduced participant burden (eg, quick, easy and convenient access, increased distance between participant and investigator allowing participants not to feel judged) [41] should help minimize attrition.

Participation in the study for altruistic reasons may be reinforced by the public nature of the research. Indeed, two-thirds of the participants considered the exclusive public funding as important or very important, with the link being particularly pronounced in men, older adults, and individuals of higher socioeconomic status. This finding is not surprising in a European context in which the majority of cohort studies are funded by public organizations. Indeed, 94% of French participants in an opinion survey conducted in the general French population reported that a large part of biomedical research needs to be funded by public funding, and 80% of responders feared that the increased participation of private funding in public research could undermine the independence of research and is damaging to certain research areas deemed less profitable [42]. In turn, the use of Internet for epidemiological studies could be viewed with suspicion by some participants because of fear that personal information might be diffused or sold to third parties or that their responses might not be anonymous [43]. In addition to reassuring participants that their personal information is kept private and safe, it seems important to point out the fact that the study is exclusively funded by public sources to investigate important public health issues.

Limitations

Our study has several limitations. First, responders were older and belonged to higher socioeconomic strata than nonrespondents, which might have led to an underestimation of the observed associations. Moreover, our result suggests that the influence of weight status on participation is open to question because participants in a long-term cohort are likely to be particularly health conscious and interested in nutritional issues. In addition, results may reflect the motives of participants accustomed to the study rather than their motives for enrollment in the Web-based cohort because the questions about motives were asked approximately 2 years after baseline. However, key information on exposure and potential confounding factors was collected during the first 2 years of participation in the cohort. Thus, a focus on the motives of those participants who actively participated 2 years after their inclusion is useful in terms of retention strategies during decisive periods of data collection in Web-based cohorts. In addition, the percentage distributions of the given reasons for participation could be biased due the use of a predefined list of response options. However, the participants had the opportunity to choose the “other” response option and to specify the exact reason for participation. Another limitation was the lack of information on reasons for declining participation because the call for participation was not delivered to a predefined list of individuals. Finally, anthropometric status was assessed by self-reporting and may have led to misclassification. However, in a separate validation study that used data from a subsample of 2513 participants, we compared self-reported and measured height and weight (and the resulting BMI) [44]. In particular, these participants had completed the anthropometric questionnaire 3 days before the clinical examination. The intraclass correlation coefficient ranged from .94 for height to .99 for weight. BMI classification was correct in 93% of the cases; the weighted kappa statistic was .89. Of 2513 participants, 23.5% were classified as overweight (BMI ≥25) with Web-based self-report versus 25.7% based on measurements by study staff, leading to a sensitivity of 88% and a specificity of 99%. For obesity, 9.1% versus 10.7% were classified as obese (BMI ≥30), respectively, leading to sensitivity of 83% and specificity of 100%.

Conclusions

Our study highlighted that the reliance on the Internet, the willingness to help advance public health research, and the exclusive public funding of the study were key motives for participation in this exclusively Web-based cohort. In addition, these motives for participation differed by sociodemographic profile and BMI, but not by lifestyle or health status. These findings can help improve retention of diverse population samples, particularly during important data collection periods.

Acknowledgments

We thank the scientists, dietitians, technicians, and assistants who helped carry out the NutriNet-Santé study, and all dedicated and conscientious volunteers. We especially thank Gwenaël Monot, Mohand Aït-Oufella, Paul Flanzy, Yasmina Chelghoum, Véronique Gourlet, Nathalie Arnault, and Laurent Bourhis. We thank Voluntis (a health care software company) and MXS (a software company specializing in dietary assessment tools) for developing the NutriNet-Santé Web-based interface according to our guidelines.

The NutriNet-Santé study is supported by the French Ministry of Health (DGS), the Institut de Veille Sanitaire (InVS), the Institut National de la Santé et de la Recherche Médicale (Inserm), the Institut National de la Recherche Agronomique (Inra), the Conservatoire National des Arts et Métiers (CNAM), the Institut National de Prévention et d’Education pour la Santé (INPES), the Fondation pour la Recherche Médicale (FRM), and the Université Paris 13.

Conflicts of Interest

None declared.

References

1. American Institute for Cancer Research / World Cancer Research Fund. Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective. Washington, DC: American Institute for Cancer Research; 2007.
2. Willett W. Nutritional Epidemiology. New York: Oxford University Press; 1998.
3. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
4. Schatzkin A, Subar AF, Moore S, Park Y, Potischman N, Thompson FE, et al. Observational epidemiologic studies of nutrition and cancer: the next generation (with better observation). Cancer Epidemiol Biomarkers Prev 2009 Apr;18(4):1026-1032 [FREE Full text] [doi: 10.1158/1055-9965.EPI-08-1129] [Medline: 19336550]

5. Internet Word Stats: Usage and Population Statistics. 2012. Internet usage statistics: the big picture URL: http://www.internetworldstats.com/stats.htm [accessed 2013-11-29] [WebCite Cache ID 6LUjnbu0m]
6. EuroStat. Internet use in households and by individuals in 2012 URL: http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/ [accessed 2013-11-29] [WebCite Cache ID 6LUjItaUk]
7. Médiamétrie. Internet 2009 URL: http://www.mediametrie.fr/internet/ [accessed 2013-11-29] [WebCite Cache ID 6LUjf7joP]
8. Smith B, Smith TC, Gray GC, Ryan MA, Millennium Cohort Study Team. When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study. Am J Epidemiol 2007 Dec 1;166(11):1345-1354 [FREE Full text] [doi: 10.1093/aje/kwm212] [Medline: 17728269]
9. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
10. Ekman A, Klint A, Dickman PW, Adami HO, Litton JE. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. Eur J Epidemiol 2007;22(5):293-300. [doi: 10.1007/s10654-006-9091-0] [Medline: 17206467]
11. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. Eur J Epidemiol 2010 May;25(5):297-304 [FREE Full text] [doi: 10.1007/s10654-010-9431-y] [Medline: 20148289]
12. Kesse-Guyot E, Andreeva V, Castetbon K, Vernay M, Touvier M, Méjean C, et al. Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Santé study. J Med Internet Res 2013;15(9):e205 [FREE Full text] [doi: 10.2196/jmir.2488] [Medline: 24036068]
13. Richiardi L, Baussano I, Vizzini L, Douwes J, Pearce N, Merletti F, NINFEA cohort. Feasibility of recruiting a birth cohort through the Internet: the experience of the NINFEA cohort. Eur J Epidemiol 2007;22(12):831-837. [doi: 10.1007/s10654-007-9194-2] [Medline: 17955333]
14. Russell CW, Boggs DA, Palmer JR, Rosenberg L. Use of a web-based questionnaire in the Black Women's Health Study. Am J Epidemiol 2010 Dec 1;172(11):1286-1291 [FREE Full text] [doi: 10.1093/aje/kwq310] [Medline: 20937635]
15. Turner C, Bain C, Schluter PJ, Yorkston E, Bogossian F, McClure R, NursesMidwives e-cohort Group. Cohort Profile: The Nurses and Midwives e-Cohort Study--a novel electronic longitudinal study. Int J Epidemiol 2009 Feb;38(1):53-60 [FREE Full text] [doi: 10.1093/ije/dym294] [Medline: 18202083]
16. West R, Gilsenan A, Coste F, Zhou X, Brouard R, Nonnemaker J, et al. The ATTEMPT cohort: a multi-national longitudinal study of predictors, patterns and consequences of smoking cessation; introduction and evaluation of internet recruitment and data collection methods. Addiction 2006 Sep;101(9):1352-1361. [doi: 10.1111/j.1360-0443.2006.01534.x] [Medline: 16911736]
17. Brouwer W, Oenema A, Raat H, Crutzen R, de Nooijer J, de Vries NK, et al. Characteristics of visitors and revisitors to an Internet-delivered computer-tailored lifestyle intervention implemented for use by the general public. Health Educ Res 2010 Aug;25(4):585-595 [FREE Full text] [doi: 10.1093/her/cyp063] [Medline: 19897515]
18. Niessen MA, Laan EL, Robroek SJ, Essink-Bot ML, Peek N, Kraaijenhagen RA, et al. Determinants of participation in a web-based health risk assessment and consequences for health promotion programs. J Med Internet Res 2013;15(8):e151 [FREE Full text] [doi: 10.2196/jmir.2387] [Medline: 23933650]
19. Robroek SJ, Lindeboom DE, Burdorf A. Initial and sustained participation in an internet-delivered long-term worksite health promotion program on physical activity and nutrition. J Med Internet Res 2012;14(2):e43 [FREE Full text] [doi: 10.2196/jmir.1788] [Medline: 22390886]
20. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [doi: 10.1093/her/cyl096] [Medline: 16971674]
21. Van 't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: a longitudinal study. J Med Internet Res 2010;12(3):e37 [FREE Full text] [doi: 10.2196/jmir.1345] [Medline: 20813716]
22. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. J Med Internet Res 2007;9(1):e1 [FREE Full text] [doi: 10.2196/jmir.9.1.e1] [Medline: 17478410]
23. Reips UD. Standards for Internet-based experimenting. Exp Psychol 2002;49(4):243-256. [Medline: 12455331]
24. Bowen DJ, Kuniyuki A, Shattuck A, Nixon DW, Sponzo RW. Results of a volunteer program to conduct dietary intervention research for women. Ann Behav Med 2000;22(1):94-100. [Medline: 10892534]
25. Hallowell N, Cooke S, Crawford G, Lucassen A, Parker M, Snowdon C. An investigation of patients' motivations for their participation in genetics-related research. J Med Ethics 2010 Jan;36(1):37-45. [doi: 10.1136/jme.2009.029264] [Medline: 20026692]
26. Kirkland SA, Raina PS, Wolfson C, Strople G, Kits O, Dukeshire S, et al. Exploring the acceptability and feasibility of conducting a large longitudinal population-based study in Canada. Can J Aging 2009 Sep;28(3):231-242. [doi: 10.1017/S0714980809990043] [Medline: 19860978]

27. Mein G, Seale C, Rice H, Johal S, Ashcroft RE, Ellison G, et al. Altruism and participation in longitudinal health research? Insights from the Whitehall II Study. Soc Sci Med 2012 Dec;75(12):2345-2352. [doi: 10.1016/j.socscimed.2012.09.006] [Medline: 23031604]
28. Williams B, Entwistle V, Haddow G, Wells M. Promoting research participation: why not advertise altruism? Soc Sci Med 2008 Apr;66(7):1451-1456. [doi: 10.1016/j.socscimed.2007.12.013] [Medline: 18222579]
29. Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. BMC Public Health 2010;10:242 [FREE Full text] [doi: 10.1186/1471-2458-10-242] [Medline: 20459807]
30. Touvier M, Méjean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. Eur J Epidemiol 2010 May;25(5):287-296. [doi: 10.1007/s10654-010-9433-9] [Medline: 20191377]
31. Touvier M, Kesse-Guyot E, Méjean C, Pollet C, Malon A, Castetbon K, et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. Br J Nutr 2011 Apr;105(7):1055-1064. [doi: 10.1017/S0007114510004617] [Medline: 21080983]
32. Vergnaud AC, Touvier M, Méjean C, Kesse-Guyot E, Pollet C, Malon A, et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Santé study. Int J Public Health 2011 Aug;56(4):407-417. [doi: 10.1007/s00038-011-0257-5] [Medline: 21538094]
33. Hercberg S, Chat-Yung S, Chaulia M. The French National Nutrition and Health Program: 2001-2006-2010. Int J Public Health 2008;53(2):68-77. [Medline: 18681335]
34. Status WECOP, Habicht J. Physical Status: The Use and Interpretation of Anthropometry. Geneva: World Health Organization; 1995.
35. Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of consent in epidemiologic research: evidence from over 25,000 responders. Am J Epidemiol 2004 Jun 1;159(11):1087-1094 [FREE Full text] [doi: 10.1093/aje/kwh141] [Medline: 15155293]
36. Eagan TM, Eide GE, Gulsvik A, Bakke PS. Nonresponse in a community cohort study: predictors and consequences for exposure-disease associations. J Clin Epidemiol 2002 Aug;55(8):775-781. [Medline: 12384191]
37. Shahar E, Folsom AR, Jackson R. The effect of nonresponse on prevalence estimates for a referent population: insights from a population-based cohort study. Atherosclerosis Risk in Communities (ARIC) Study Investigators. Ann Epidemiol 1996 Nov;6(6):498-506. [Medline: 8978880]
38. Cunradi CB, Moore R, Killoran M, Ames G. Survey nonresponse bias among young adults: the role of alcohol, tobacco, and drugs. Subst Use Misuse 2005;40(2):171-185. [Medline: 15770883]
39. Wild TC, Cunningham J, Adlaf E. Nonresponse in a follow-up to a representative telephone survey of adult drinkers. J Stud Alcohol 2001 Mar;62(2):257-261. [Medline: 11327192]
40. Corbie-Smith GM. Minority recruitment and participation in health research. N C Med J 2004;65(6):385-387. [Medline: 15714732]
41. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
42. IPSOS Survey. Forum Science, Recherche et Société, organized by Le Monde et La Recherche URL: http://www.axa-research.org/fr/forum-sciences-recherche-et-societe-au-college-de-france [accessed 2014-06-25] [WebCite Cache ID 6QamsQva6]
43. Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999 Aug;31(3):433-438. [Medline: 10502866]
44. Lassale C, Péneau S, Touvier M, Julia C, Galan P, Hercberg S, et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Santé study. J Med Internet Res 2013;15(8):e152 [FREE Full text] [doi: 10.2196/jmir.2575] [Medline: 23928492]

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 10.12.13; peer-reviewed by A Burdorf, B Lohse; comments to author 23.04.14; revised version received 25.06.14; accepted 22.07.14; published 07.08.14.

Please cite as:
Méjean C, Szabo de Edelenyi F, Touvier M, Kesse-Guyot E, Julia C, Andreeva VA, Hercberg S
Motives for Participating in a Web-Based Nutrition Cohort According to Sociodemographic, Lifestyle, and Health Characteristics: The NutriNet-Santé Cohort Study
J Med Internet Res 2014;16(8):e189
URL: http://www.jmir.org/2014/8/e189/
doi:10.2196/jmir.3161
PMID:25135800

Original Paper

Numeracy and Literacy Independently Predict Patients' Ability to Identify Out-of-Range Test Results

Brian J Zikmund-Fisher[1,2,3,4], PhD; Nicole L Exe[3], MPH; Holly O Witteman[5,6,7], PhD

[1]Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI, United States
[2]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States
[3]Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI, United States
[4]Risk Science Center, University of Michigan, Ann Arbor, MI, United States
[5]Department of Family and Emergency Medicine, Faculty of Medicine, Université Laval, Québec City, QC, Canada
[6]Office of Education and Continuing Professional Development, Faculty of Medicine, Université Laval, Québec City, QC, Canada
[7]Research Centre of the CHU de Québec, Québec City, QC, Canada

Corresponding Author:
Brian J Zikmund-Fisher, PhD
Department of Health Behavior and Health Education
University of Michigan
1415 Washington Heights
Ann Arbor, MI, 48109
United States
Phone: 1 7349369179
Fax: 1 7347637379
Email: bzikmund@umich.edu

Abstract

Background: Increasing numbers of patients have direct access to laboratory test results outside of clinical consultations. This offers increased opportunities for both self-management of chronic conditions and advance preparation for clinic visits if patients are able to identify test results that are outside the reference ranges.

Objective: Our objective was to assess whether adults can identify laboratory blood test values outside reference ranges when presented in a format similar to some current patient portals implemented within electronic health record (EHR) systems.

Methods: In an Internet-administered survey, adults aged 40-70 years, approximately half with diabetes, were asked to imagine that they had type 2 diabetes. They were shown laboratory test results displayed in a standard tabular format. We randomized hemoglobin A_{1c} values to be slightly (7.1%) or moderately (8.4%) outside the reference range and randomized other test results to be within or outside their reference ranges (ie, multiple deviations). We assessed (1) whether respondents identified the hemoglobin A_{1c} level as outside the reference range, (2) how respondents rated glycemic control, and (3) whether they would call their doctor. We also measured numeracy and health literacy.

Results: Among the 1817 adult participants, viewing test results with multiple deviations increased the probability of identifying hemoglobin A_{1c} values as outside the reference range (participants with diabetes: OR 1.47, 95% CI 1.12-1.92, P=.005; participants without diabetes: OR 1.50, 95% CI 1.13-2.00, P=.005). Both numeracy and health literacy were significant predictors of correctly identifying out-of-range values. For participants with diabetes, numeracy OR 1.32 per unit on a 1-6 scale (95% CI 1.15-1.51, P<.001) and literacy OR 1.59 per unit of a 1-5 scale (95% CI 1.35-1.87, P<.001); for participants without diabetes, numeracy OR 1.36 per unit (95% CI 1.17-1.58, P<.001) and literacy OR 1.33 per unit (95% CI 1.12-1.58, P=.001). Predicted probabilities suggested 77% of higher numeracy and health literacy participants, but only 38% of lower numeracy and literacy participants, could correctly identify the hemoglobin A_{1c} levels as outside the reference range. Correct identification reduced perceived blood glucose control (mean difference 1.68-1.71 points on a 0-10 scale, P<.001). For participants with diabetes, increased health literacy reduced the likelihood of calling one's doctor when hemoglobin A_{1c}=7.1% (OR 0.66 per unit, 95% CI 0.52-0.82, P<.001) and increased numeracy increased intention to call when hemoglobin A_{1c}=8.4% (OR 1.36 per unit, 95% CI 1.10-1.69, P=.005).

Conclusions: Limited health literacy and numeracy skills are significant barriers to basic use of laboratory test result data as currently presented in some EHR portals. Regarding contacting their doctor, less numerate and literate participants with diabetes appear insensitive to the hemoglobin A_{1c} level shown, whereas highly numerate and literate participants with diabetes appear very sensitive. Alternate approaches appear necessary to make test results more meaningful.

(J Med Internet Res 2014;16(8):e187) doi:10.2196/jmir.3241

KEYWORDS

numeracy; literacy; patient education as topic; electronic health records

Introduction

Increasing numbers of patients have direct access to laboratory test results outside of clinical consultations via patient portals implemented within electronic health record (EHR) systems. Patient access to such records was included in the Centers for Medicare and Medicaid Services (CMS) Meaningful Use of Health Information Technology criteria from 2010 [1], and already exists in some large health care systems. Strong federal incentives supporting adoption of electronic medical record systems will likely significantly increase the availability of patient portals in the future.

Patients use such systems to view medical test results and value being able to do so [2-4]. Direct patient use of test data is consistent with trends toward patient-centered approaches to care, patient engagement, and the medical home concept, all of which encourage greater patient involvement in both medical decision making and health self-management [5-7]. In that manner, patient access to such health data promotes a transfer of some of the responsibility for health management from care providers to the patients themselves [8]. Patient access is also congruent with the trend for people to actively gather, manage, and analyze their personal data (eg, the “quantified self” movement) [9]. Perhaps most importantly, there is an ethical imperative to provide easy access to patients who want it [10].

Patients want to be notified of laboratory test results, regardless of whether the findings were normal or abnormal [11], because failures to inform patients of test results are unfortunately all too common, even for abnormal or otherwise actionable test results [12]. Direct access enables patients to seek out their results by themselves, thereby providing a second opportunity for identifying actionable results and preventing unnecessary harm.

Test result data can also enable patients to better prepare for clinic visits by focusing their attention on test results that are abnormal or of concern. This knowledge could lead patients to prepare questions or seek out relevant information before the visit. Such preparation benefits patients, but it also benefits the health care system by making visits more efficient [1].

Patients can also use test results to improve self-management of their current health conditions [1,13]. For example, a person with diabetes could both assess her current status and identify long-term trends in her blood glucose control. She could use such data to determine whether her current health management efforts (eg, behavior programs, medications) are working. Such information offers the potential to increase patient activation and the likelihood of engaging in particular treatment or health behaviors [1].

Achievement of these potential benefits, however, requires patients to perform a simple, yet critical, task: to be able to correctly identify which test results are out of range (ie, outside the reference ranges) from the (usually) much larger set of data provided. Unfortunately, there are several reasons to suspect that many patients will have difficulty with this task when the test results are displayed in the tabular format currently used in many interface designs.

First, many patients have limited health literacy, which inhibits their ability to interpret the health information they read and use that information to manage their health [14-16]. For example, low health literacy is associated with less knowledge about the medications one is taking [17], being less able to read and understand medication labels [18], unintentional nonadherence to hospital discharge instructions [19], and increased mortality [20,21]. Health literacy affects patient use of laboratory test results in 2 ways. First, lack of attention to issues of health literacy when designing patient portals limits the accessibility of such tools and, therefore, limits their impact among those who might most benefit from them. Even restricting analysis to people with Internet access, lower health literacy is associated with a lower likelihood of logging into a patient portal in the first place [22]. Indeed, patient portal use is lower among the more vulnerable populations [23]. Second, health literacy affects patients’ abilities to gather background and contextual information (eg, about what a test is, what values are normal or concerning) necessary to cognitively evaluate the meaning of a test result in relation to their health.

Second, many patients also have lower numeracy skills (ie, poor ability to use and draw meaning from numbers) [24,25]. Although some measures of health literacy (eg, Test of Functional Health Literacy in Adults, TOFHLA) include assessments of what is variably termed numerical ability or quantitative literacy, there is growing evidence that numeracy is a distinct construct that is particularly relevant to data interpretation tasks. Numeracy predicts people’s ability to read nutrition labels, calculate medication dosages, maintain anticoagulation control, and maintain glycemic control better than measures of health literacy do [26-29]. Patients with lower numeracy skills may lack the capacity to interpret test outcome data in some current presentations. In addition, numeracy appears necessary for people to develop emotional responses to data [30]. This is problematic given the large amount of theoretical and experimental evidence that emotions are both integral to risk perceptions and necessary for effective decision making [31-34]. As a result, less numerate patients are unlikely to know how to use medical test results if they cannot get a feeling of “goodness” or “badness” from the data [30].

Without careful design that attends to issues of health literacy and numeracy, presentations of laboratory test results (whether in patient portals or via a clinician's office) could be of little use to less literate and numerate patients. Although some initiatives have used cues such as color to help patients identify out-of-range values [35], often laboratory results are shown in the same tabular format that is provided to clinicians. These tables present a dozen or more tests simultaneously, usually labeled with unfamiliar abbreviations, reported in unfamiliar units, and lacking guidance as to whether higher numbers represent more positive or negative outcomes. Unfortunately, less numerate people have particular difficulty identifying decision-relevant information out of larger sets of data [36]. Therefore, the sheer volume of information available through patient portals is particularly challenging for the less numerate [8].

We designed an experimental study to assess the degree that adults, especially those with lower numeracy and/or lower health literacy, are able or not able to identify out-of-range values in prototypical medical test result displays. Participants viewed multiple panels of test results typical of what would be ordered for ongoing management of a person with type 2 diabetes and were asked to (1) identify all values outside the reference range, (2) assess the degree of blood glucose control represented by those results, and (3) identify whether they would call their doctor regarding these results. To test patient sensitivity to variations in test results, we experimentally varied 2 factors: (1) hemoglobin A_{1c} levels were mildly or moderately elevated and (2) other test results were within or outside their reference ranges. To enable assessment of the role of numeracy and health literacy skills on people's ability to complete these tasks accurately, all participants completed validated measures of both constructs.

Methods

Participants

We recruited a stratified random sample of US adults aged 40-70 years from a panel of Internet users administered by Survey Sampling International (SSI, Shelton, CT, USA), which recruits panel members through various opt-in methods. To ensure demographic diversity (although not representativeness) and offset variations in response rates, we drew subsamples by both age and race (thereby approximating the distributions of these characteristics in the US population). We also drew separate subsamples by experience with diabetes: We specified that approximately half of completed surveys be from panel participants who had previously indicated that they had diabetes (and hence might have had greater knowledge about hemoglobin A_{1c} tests) and half from people without personal experience with diabetes (who might be more similar to newly diagnosed patients). The number of email invitations in each subsample was dynamically adjusted until quotas were achieved.

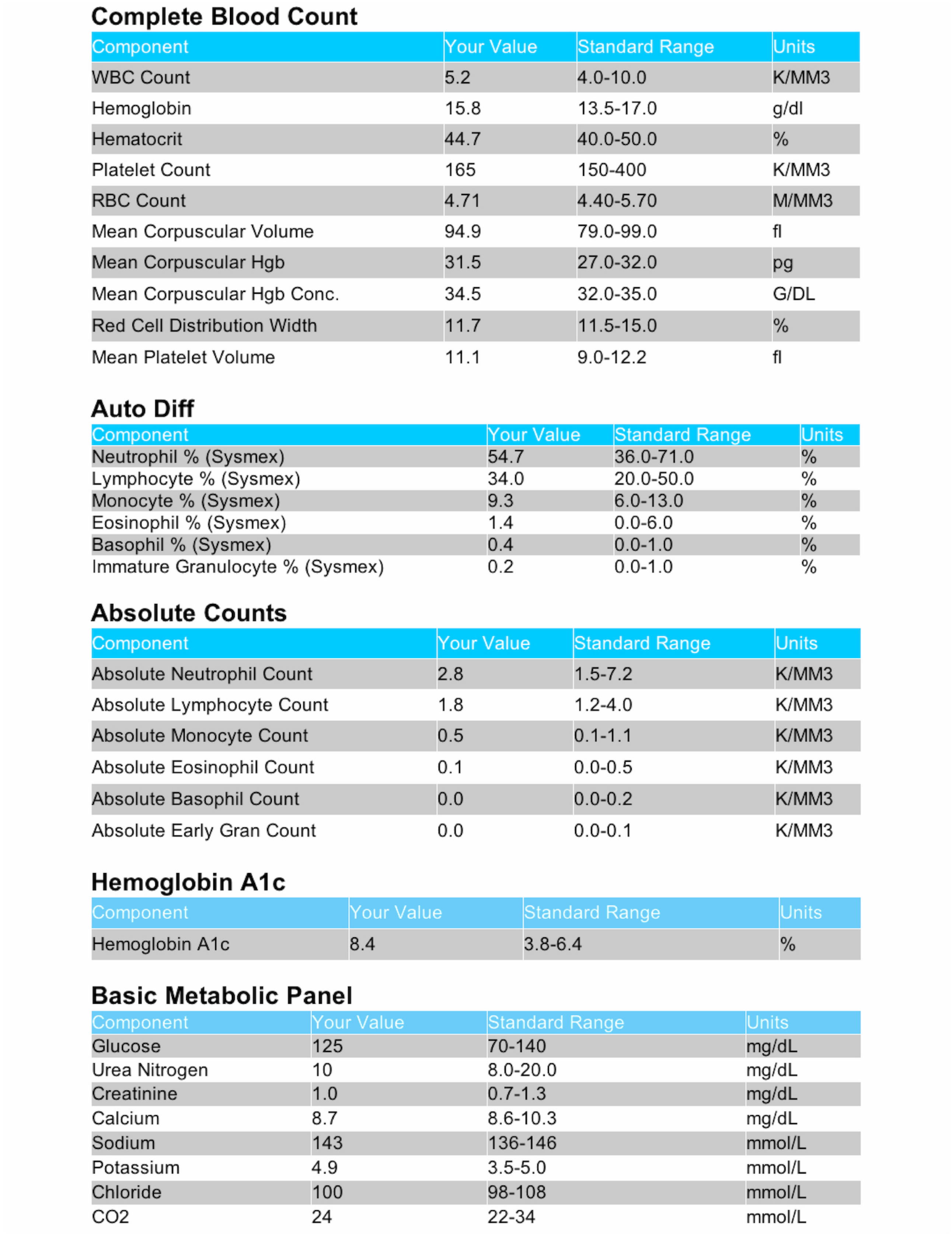
Selected panel members received email invitations with a personalized link (tracked to prevent duplicates) and nonresponders received 1 reminder email. Those who clicked on the link then viewed an introductory page that provided information about the estimated length of the survey (10 to 15 minutes), the purpose of the study, and affiliation and contact information for the investigators before taking the participant to the main study materials. We recruited for a 2-week period in January 2013. On completion, participants were entered into instant-win contests and regular draws administered by SSI for modest prizes.

Design

Participants were asked to imagine that they were diagnosed with type 2 diabetes, had been maintaining good blood glucose control with a previous hemoglobin A_{1c} test result of 6.8%, and had an explicit goal of maintaining hemoglobin A_{1c} values below 7%. Participants were then asked to imagine that they were viewing the results of a set of blood tests (complete blood cell count, CBC; hemoglobin A_{1c}; and renal panel) that had been ordered between doctor's visits. Following the format currently implemented in the patient portal of a major academic medical center, all tables showed test values, standard ranges, and units, but did not show indicators for high or low values (the medical center includes high/low indicators in clinician interfaces but omits them from the patient interface). As shown in Figure 1, all tests were presented on a single page grouped by panel per standard practice.

We manipulated the test results shown in a 2×2 factorial design. All participants viewed results that showed that hemoglobin A_{1c} was elevated above the standard range (reported as 3.8%-6.4%). We randomly varied the degree of A_{1c} elevation by randomizing participants to view a hemoglobin A_{1c} result of either 7.1% or 8.4%. Thus, both values should be identified as out of range, but only the 8.4% value is sufficiently high (and a large enough change from the previous value) to potentially warrant more timely attention. In addition, we independently varied whether all other reported results were within standard ranges (single deviation condition) or whether multiple results were out of range (multiple deviations condition). Participants in the multiple deviations condition saw tables with out-of-range values for white blood cell (WBC) count, platelet count, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), neutrophil %, lymphocyte %, monocyte %, absolute neutrophil count, and serum glucose. These values were either elevated or reduced to be consistent with a temporary viral infection.

Figure 1. Screenshot of the test results display (hemoglobin A1c=8.4%, single deviation condition).



Complete Blood Count

Component	Your Value	Standard Range	Units
WBC Count	5.2	4.0-10.0	K/MM3
Hemoglobin	15.8	13.5-17.0	g/dl
Hematocrit	44.7	40.0-50.0	%
Platelet Count	165	150-400	K/MM3
RBC Count	4.71	4.40-5.70	M/MM3
Mean Corpuscular Volume	94.9	79.0-99.0	fl
Mean Corpuscular Hgb	31.5	27.0-32.0	pg
Mean Corpuscular Hgb Conc.	34.5	32.0-35.0	G/DL
Red Cell Distribution Width	11.7	11.5-15.0	%
Mean Platelet Volume	11.1	9.0-12.2	fl

Auto Diff

Component	Your Value	Standard Range	Units
Neutrophil % (Sysmex)	54.7	36.0-71.0	%
Lymphocyte % (Sysmex)	34.0	20.0-50.0	%
Monocyte % (Sysmex)	9.3	6.0-13.0	%
Eosinophil % (Sysmex)	1.4	0.0-6.0	%
Basophil % (Sysmex)	0.4	0.0-1.0	%
Immature Granulocyte % (Sysmex)	0.2	0.0-1.0	%

Absolute Counts

Component	Your Value	Standard Range	Units
Absolute Neutrophil Count	2.8	1.5-7.2	K/MM3
Absolute Lymphocyte Count	1.8	1.2-4.0	K/MM3
Absolute Monocyte Count	0.5	0.1-1.1	K/MM3
Absolute Eosinophil Count	0.1	0.0-0.5	K/MM3
Absolute Basophil Count	0.0	0.0-0.2	K/MM3
Absolute Early Gran Count	0.0	0.0-0.1	K/MM3

Hemoglobin A1c

Component	Your Value	Standard Range	Units
Hemoglobin A1c	8.4	3.8-6.4	%

Basic Metabolic Panel

Component	Your Value	Standard Range	Units
Glucose	125	70-140	mg/dL
Urea Nitrogen	10	8.0-20.0	mg/dL
Creatinine	1.0	0.7-1.3	mg/dL
Calcium	8.7	8.6-10.3	mg/dL
Sodium	143	136-146	mmol/L
Potassium	4.9	3.5-5.0	mmol/L
Chloride	100	98-108	mmol/L
CO2	24	22-34	mmol/L

Outcome Measures

We asked participants to answer a series of questions about the test results display, which remained visible so that the questions would measure test understanding and interpretation, not recall.

Participants first rated their perceived blood glucose control based on the hypothetical scenario and data by answering the question "according to these test results, how well have you been keeping your blood sugars under control *over the last 3* *months*?” (emphasis in original), with responses on an 11-point scale ranging from not at all well to extremely well.

The next question asked participants to “please mark which tests had results different than what they should be” (ie, had out-of-range values) using a set of checkboxes. If participants marked any tests as being out of range, they received a follow-up question asking them to “please rate how concerned you would be about each of these identified tests” on a 5-point scale from not at all concerned to extremely concerned. Our analysis focused on whether participants identified the hemoglobin A_{1c} value as out of range and their subsequent ratings of concern.

Next, we asked participants to indicate what they would do after reviewing the test results (with test results no longer visible). Participants chose from 3 options: (1) call or email doctor’s office and ask to speak with your doctor immediately, (2) call or email doctor’s office to see whether you can schedule an appointment with your doctor in the next few weeks (your next scheduled appointment is currently in 3 months), or (3) talk to your doctor about these results at your next appointment in 3 months.

We also asked 2 questions to measure perceived usefulness. Participants rated “how well did you understand what the test results said” on a 5-point scale from did not understand at all to understood completely and “how useful were these test results” on a 5-point scale from not at all useful to extremely useful.

Individual Difference Measures

Participants next completed several individual difference measures that we hypothesized might interact with ability to interpret test result tables. Because ample evidence exists that even highly educated adults can have poor numeracy skills (ie, facility and comfort with quantitative health information such as risk statistics) [25,37,38], all study participants completed the Subjective Numeracy Scale (SNS) [39]. The SNS measures both perceived quantitative ability and preference for receiving information in numerical form and has previously been shown to correlate with the ability to recall and comprehend both textual and graphical risk communications [40,41]. A participant’s SNS score is calculated as his or her mean rating across the 8 SNS questions (after accounting for reverse coding) and ranges from 1 (least numerate) to 6 (most numerate). In addition, participants also completed Chew et al’s [42] 3-question measure of limited health literacy, which has been validated and shown to be highly correlated with other measures, such as the Rapid Estimate of Adult Literacy in Medicine (REALM) and Short Test of Functional Health Literacy in Adults (S-TOFHLA). Participants’ literacy score was the mean response for the 3 questions (after reverse coding 1 question) and ranges from 1 (least literate) to 5 (most literate).

Participants also completed standard demographic questions and indicated whether they were diagnosed previously with diabetes. This latter direct-response measure was used for analysis of the effect of diabetes experience instead of the information from SSI that had been used to guide the sampling process.

Data Management

All data were collected anonymously using the Qualtrics online survey platform. Participants were identified and prevented from taking the survey multiple times via unique identification numbers provided by SSI within the redirected URL. The design, sampling process, data management procedures, and outcome measures received exempt status approval from the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board.

Statistical Analyses

We performed separate analyses of data from participants who self-reported that they were diagnosed with diabetes and from nondiabetic participants. For each group, we conducted chi-square and logistic regression analyses of whether respondents identified the hemoglobin A1c value as out of range and whether they intended to call their doctor about the test results. We conducted *t* tests, correlation analyses, and linear regression analyses of both perceived blood glucose control and perceived usefulness of the test result displays. Regression analyses included indicator variables for experimental factors, education, and respondents’ health literacy and numeracy scale scores as continuous variables. Education was modeled as 2 indicator variables for (1) greater than high school education of some type, but no bachelor degree and (2) a bachelor or higher degree (each compared to a baseline group of high school education or lower). We also conducted additional analyses to determine whether the ability to correctly identify a test result as out of range mediated any effects of experimental or demographic predictors on perceptions of blood glucose control and intentions to call a doctor. All analyses were performed using Stata 12 (StataCorp LP, College Station, TX, USA), and all tests of significance were 2-sided and used alpha=.05.

Results

Participant Characteristics

In total, 1817 people aged between 40 and 70 years completed the survey. Participant characteristics are reported in Table 1. We observed a wide range of educational achievement with 31.36% (567/1808) of participants having a bachelor or higher college degree, but also 23.73% (429/1808) with an education level of high school or less.

Within our sample, the SNS numeracy measure showed high reliability (Cronbach alpha=.87), and the mean SNS score was 4.47 (SD 1.06, range 1.0-6.0). Mean score on the health literacy measure was 3.84 (SD 0.87, range 1-5), although the scale showed relatively weak reliability (Cronbach alpha=.54). These 2 measures were moderately correlated (r=.26), although 189 of 1799 (10.51%) participants indicated lower numeracy (SNS $\leq$4) and higher literacy (literacy $\geq$4) and 117 of 1799 (6.50%) participants had the reverse pattern of higher numeracy (SNS $\geq$5) and lower literacy (literacy $\leq$3).

Table 1. Participant characteristics (N=1817).

Characteristic and categories	n (%)[a]	Mean (SD)
Age (years) (n=1814)		54.2 (8.4)
40-49	635 (35.01)	
50-59	605 (33.35)	
60-70	574 (31.64)	
Sex (n=1814)		
Male	901 (49.67)	
Female	913 (50.33)	
Ethnicity (n=1807)		
Hispanic (any race)	170 (9.41)	
Race [b] (n=1810)		
White	1407 (77.73)	
African-American	280 (15.47)	
All other	161 (8.90)	
Education (n=1808)		
≤High school	429 (23.73)	
Some college/trade	812 (44.91)	
Bachelor/master/doctorate degree	567 (31.36)	
Subjective Numeracy Scale Score (n=1804)		4.47 (1.06)
1.00-1.99	42 (2.33)	
2.00-2.99	138 (7.65)	
3.00-3.99	315 (17.46)	
4.00-4.99	600 (33.26)	
5.00-5.99	644 (35.70)	
6.00	65 (3.60)	
Limited Health Literacy Scale Score (n=1799)		3.84 (0.87)
1.00-1.99	19 (1.06)	
2.00-2.99	219 (12.17)	
3.00-3.99	649 (36.08)	
4.00-4.99	576 (32.02)	
5.00	336 (18.68)	
Participant with diabetes (n=1812)	971 (53.59)	

[a] Reports results only for those respondents who completed each question or measure.

[b] Respondents could indicate more than 1 race.

Identification of Hemoglobin A1c Value as Out of Range

Overall, approximately half (931/1817, 51.24%) of participants correctly identified the hemoglobin A_{1c} value as being “different than what [it] should be.” Participants with diabetes were more likely to identify the out-of-range hemoglobin A_{1c} value than participants without diabetes were (participants with diabetes: 546/971, 56.2%; participants without diabetes: 384/841, 45.7%, χ^2_1=20.2, P<.001). Rates of correctly identifying out-of-range hemoglobin A_{1c} values were also significantly higher among participants in the multiple deviations condition versus those in the single deviation condition (multiple deviations: 499/898, 55.6%; single deviation: 432/919, 47.0%, χ^2_1=13.3, P<.001). The specific hemoglobin A_{1c} value reported had no effect on the likelihood of marking it as out of range (hemoglobin A_{1c}=7.1%: 462/911, 50.7%; hemoglobin A_{1c}=8.4%: 469/906, 51.8%; χ^2_1=0.2, P=.65).

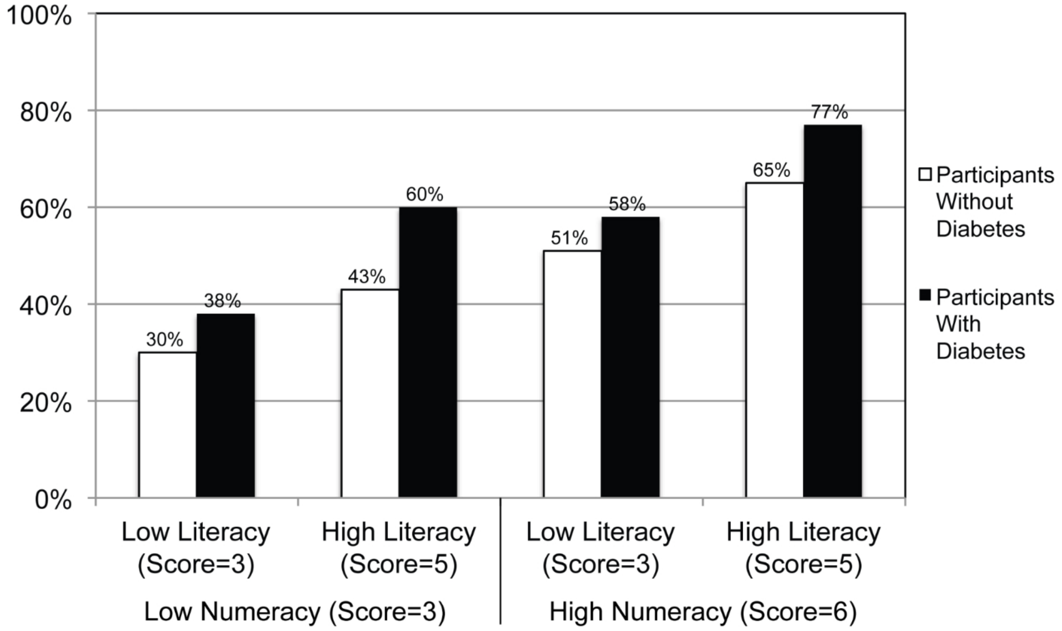
Table 2 reports logistic regression analyses identifying predictors of correctly identifying hemoglobin A_{1c} values as out of range for participants with and without diabetes. The multivariate analysis confirms the significant effect of the multiple deviations condition for both groups of participants. However, the specific hemoglobin A_{1c} value shown was a significant predictor of correctly identifying it as out of range for participants with diabetes (more likely to mark if hemoglobin A_{1c}=8.4% vs 7.1%). In addition, the regression analyses identified significant and independent effects of both participant numeracy and health literacy as well as weaker effects of education.

Table 2. Logistic regression results showing predictors of identifying hemoglobin A_{1c} levels as out of range for participants with and without diabetes.

Variable	Participants without diabetes (n=827)			Participants with diabetes (n=963)		
	OR	95% CI	*P*	OR	95% CI	*P*
Hemoglobin A_{1c} test result=8.4% (vs=7.1%)	0.85	0.64, 1.13	.26	1.38	1.05, 1.80	.02
Multiple deviations condition (vs single)	1.50	1.13, 2.00	.005	1.47	1.12, 1.92	.005
Education: high school or less	—	—	—	—	—	—
Education: >high school but <bachelor degree	1.28	0.88, 1.86	.20	1.17	0.83, 1.66	.36
Education: bachelor degree or higher	1.41	0.94, 2.11	.10	1.53	1.03, 2.27	.04
Subjective numeracy score (per unit, range 1-6)	1.36	1.17, 1.57	<.001	1.32	1.15, 1.51	<.001
Literacy score (per unit, range 1-5)	1.33	1.12, 1.58	.001	1.59	1.35, 1.87	<.001

To clarify the effect size of the experience with diabetes, numeracy, and health literacy effects, we calculated the predicted likelihood that participants with different combinations of lower versus higher numeracy and lower versus higher health literacy would mark hemoglobin A_{1c} levels as out of range, holding all other predictors to their mean values. We conducted this analysis separately for participants with and without diabetes. Although we recognize that numeracy and literacy are often at least moderately correlated in practice (in our sample: *r*=.26), the predicted probabilities help to clarify the independent and combined effects of these 2 factors over the range of possible patient skill levels. We defined lower numeracy or health literacy as a score of 3 on these scales (corresponding to the tenth to thirteenth percentile of the observed distribution for each measure) and higher numeracy as the scale maximums of 6 for numeracy and 5 for health literacy. As Figure 2 shows, the combined effect of lower health literacy and lower numeracy more than halves the probability of identifying out-of-range values (from 77% to 38% for participants with diabetes and from 65% to 30% for participants without diabetes).

Figure 2. Predicted probabilities that participants with and without diabetes would correctly identify hemoglobin A1c test results as outside the standard range by lower versus higher literacy and numeracy levels.



Perceptions of Blood Glucose Control

Perceptions of blood glucose control averaged in the middle of the 0-10 scale, but we observed substantial variance. When the displayed hemoglobin A_{1c} level was 7.1%, the mean perception of blood glucose control was 5.62 (SD 2.87) and did not vary significantly by diabetes diagnosis. However, when the displayed hemoglobin A1c level was 8.4%, not only were perceptions of control lower (mean 4.71, SD 3.17) but perceptions of the participants with diabetes were significantly lower than those of participants without diabetes (participants with diabetes: mean 4.40, SD 3.13; participants without diabetes: mean 5.05, SD 3.18; t_{887}=3.07, *P*=.002).

Table 3 reports linear regressions predicting perceived blood glucose control among participants with and without diabetes. Because the act of identifying the hemoglobin A_{1c} test result as being outside the standard range is a likely precursor of perceiving one's blood glucose as being less controlled, we included that variable as an independent predictor in addition to the previous model predictors.

Confirming our expectations, correctly identifying hemoglobin A_{1c} levels as out of range had a highly significant and large effect on participants' ratings of their blood glucose control, reducing ratings of blood glucose control by 1.68-1.71 points on the 11-point scale (*P*<.001 for both). Also, viewing a hemoglobin A_{1c} test result of 8.4% instead of 7.1% also lowered perceptions of blood glucose control (*P*<.001 for both).

However, we also observed independent effects of health literacy and numeracy for participants without diabetes. Having higher health literacy decreased perceived glucose control by 0.54 points per unit on the 5-point literacy scale (*P*<.001), whereas higher numeracy increased perceived glucose control by 0.27 points per unit on the 6-point numeracy scale (*P*=.01).

The fact that health literacy and numeracy predicted correctly identifying the hemoglobin A_{1c} value as out of range raises the possibility that their effects on perceived blood glucose control might be partially mediated through that action. Yet, reduced models omitting the "marked A_{1c} as out-of-range" variable (not shown) were similar to those shown in Table 3. Among participants without diabetes, test result (hemoglobin A_{1c} value of 8.4% vs 7.1%) and health literacy remained highly significant, and numeracy was actually less significant (beta=0.15, *P*=.16). Among participants with diabetes, the effects of test result and high education remained highly significant, and the coefficient for health literacy became significant (beta=–0.35, *P*=.002). This suggests that correctly identifying hemoglobin A_{1c} levels as out of range may have partially mediated the effect of health literacy among participants with diabetes, but we saw no evidence of any mediation effects among participants without diabetes.

Table 3. Linear regression results showing predictors of perceived blood sugar control among participants with and without diabetes.

Variable	Participants without diabetes (n=820)		Participants with diabetes (n=953)	
	Coefficient	*P*	Coefficient	*P*
Hemoglobin A_{1c} test result=8.4% (vs 7.1%)	–0.65	.001	–1.08	<.001
Multiple alert condition (vs single alert)	–0.37	.07	–0.13	.46
Education: high school or less	—	—	—	—
Education: >high school but <bachelor degree	0.24	.37	–0.54	.03
Education: bachelor's degree or higher	–0.35	.23	–0.76	.005
Subjective numeracy score (per unit, range 1-6)	0.27	.01	0.09	.33
Literacy score (per unit, range 1-5)	–0.54	<.001	–0.17	.11
Marked hemoglobin A_{1c} result as out of range	–1.68	<.001	–1.71	<.001
Constant	7.52		7.29	

Behavioral Intentions Regarding Contacting a Doctor

Table 4 reports the proportion of respondents in each experimental condition who indicated that they would intend to call their doctor to discuss the laboratory test results, organized by participants with diabetes and those without. Most participants would call their doctor in all conditions (1218/1765, 69.01%), but in the single deviation condition, the intention to call was significantly lower when hemoglobin A_{1c} level was 7.1% vs 8.4% (A_{1c}=7.1%: χ^2_1=15.4, *P*<.001; A_{1c}=8.4%: χ^2_1=18.9, *P*<.001). However, in the multiple deviation condition, the difference in rates was significant (but only barely so) among participants with diabetes (χ^2_1=4.3, *P*=.04) and was nonsignificant among participants without diabetes (χ^2_1=0.1, *P*=.76).

Table 4. Proportion of respondents indicating they would call their doctor to discuss the laboratory test results (either immediately or rebook a set appointment to an earlier date/time) by diabetes diagnosis and experimental condition (N=1763).

Condition	Participants without diabetes (n=813)		Participants with diabetes (n=950)	
	A_{1c}=7.1%	A_{1c}=8.4%	A_{1c}=7.1%	A_{1c}=8.4%
Single deviation condition, n/n (%)	137/215 (63.7%)	178/221 (80.5%)	119/232 (51.3%)	160/225 (71.1%)
Multiple deviation condition, n/n (%)	138/186 (74.2%)	139/191 (72.8%)	167/253 (66.0%)	179/240 (74.6%)

We next report the results of logistic regression analyses of intent to call a doctor for participants with and without diabetes (Table 5). Within each table, we report separate analyses for the hemoglobin A_{1c}=7.1% and 8.4% scenarios because they represent 2 distinct scenarios that should logically evoke different behaviors in participants and therefore might have fundamentally different predictors.

Table 5. Logistic regression results showing predictors of intent to call a doctor among study participants with and without diabetes.

Variable	Participants without diabetes						Participants with diabetes					
	A_{1c}=7.1%			A_{1c}=8.4%			A_{1c}=7.1%			A_{1c}=8.4%		
	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
Multiple deviations condition (vs single)	1.47	0.94, 2.29	.09	0.54	0.33, 0.88	.01	1.79	1.23, 2.62	.003	1.11	0.73, 1.70	.63
Education: high school or less	—	—	—	—	—	—	—	—	—	—	—	—
Education: >high school but <bachelor degree	0.81	0.46, 1.43	.47	1.45	0.78, 2.69	.24	1.08	0.66, 1.76	.77	0.85	0.49, 1.50	.58
Education: bachelor degree or higher	0.98	0.51, 1.87	.95	0.95	0.50, 1.81	.88	1.36	0.79, 2.36	.27	0.64	0.33, 1.25	.19
Subjective numeracy score (per unit, range 1-6)	1.09	0.87, 1.35	.46	1.17	0.92, 1.48	.20	0.94	0.77, 1.15	.53	1.36	1.10, 1.69	.005
Literacy score (per unit, range 1-5)	0.86	0.65, 1.12	.26	0.91	0.67, 1.23	.54	0.66	0.52, 0.82	<.001	0.92	0.71, 1.20	.55
Marked A_{1c} as out of range	1.98	1.26, 3.11	.003	3.28	1.91, 5.61	<.001	1.95	1.31, 2.89	.001	2.31	1.48, 3.61	<.001

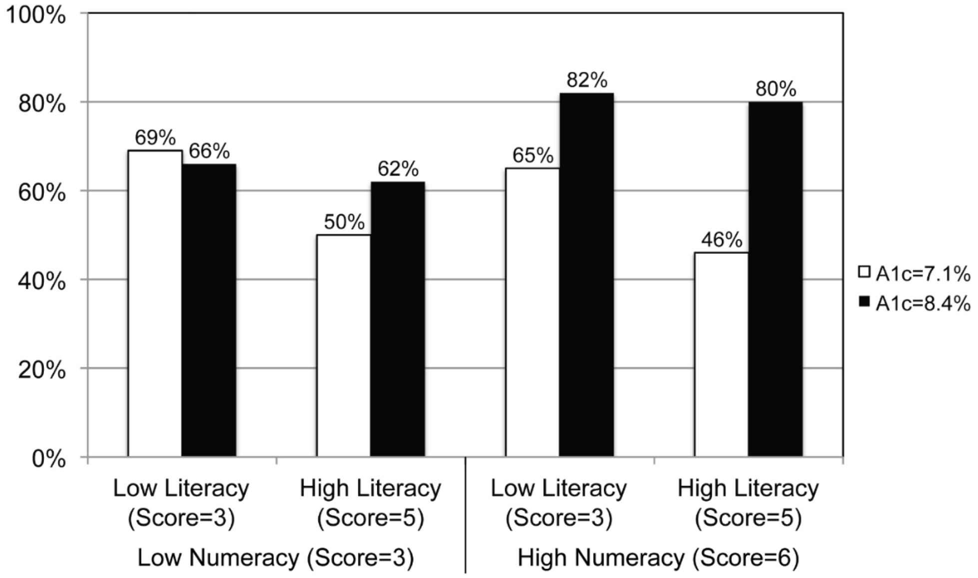
Among participants without diabetes, correctly identifying hemoglobin A_{1c} levels as out of range was the primary predictor of whether the participant intended to call their doctor. However, we also saw an interesting pattern regarding the multiple deviations condition. Having other test results (beyond hemoglobin A_{1c}) out of range tended (nonsignificantly) to increase the odds of calling the doctor if the hemoglobin A_{1c} level was 7.1% but significantly lowered the likelihood of calling the doctor if the hemoglobin A_{1c} level was 8.4%. It is unclear whether the latter effect reflects beliefs that the elevated hemoglobin A_{1c} level is not as concerning in the presence of the other nonnormal test results or simple confusion or distraction. In either case, neither health literacy nor subjective numeracy scores predicted intentions to call the doctor's office among participants without diabetes. Omitting the "marked A_{1c} as out of range" variable had little effect on either regression. The only change of note was that, in the hemoglobin A_{1c} level equals 7.1% condition, the odds ratio for the multiple deviations condition increased slightly and became statistically significant (OR 1.59, 95% CI 1.03-2.45, *P*=.04).

Among study participants with diabetes, we saw a distinct pattern of results related to health literacy and numeracy even after controlling for the continued large and significant effect of having correctly identified hemoglobin A_{1c} levels as out of range (which was highly predicted by health literacy and numeracy). Among participants with diabetes presented with hemoglobin A_{1c}=7.1% test results, increased health literacy significantly reduced the likelihood of calling one's doctor. This is consistent with these individuals having absorbed the background knowledge provided in the scenario sufficiently to recognize that a 7.1% value is only mildly elevated in comparison to the previous value of 6.8% that was provided in the scenario and was very close to the 7% threshold that was stated explicitly as the patient's goal level. However, numeracy skills were not associated with intentions to call the doctor.

Conversely, among participants with diabetes who were presented with hemoglobin A_{1c}=8.4% test results (a value that is both significantly elevated on an absolute level and a much larger increase in reference to the previous test result cited in the scenario), health literacy had no effect on intentions to call their doctor. Instead, among participants with diabetes, it was increased numeracy skills that significantly increased their intentions to call the doctor's office. This finding is consistent with the hypothesis that numeracy skills were associated with respondents' ability to recognize not merely that the 8.4% value was out of range, but that the increase of less than 2 absolute percentage points nonetheless represented a substantial and concerning change worthy of action. Both this effect and the effect of health literacy when the A_{1c} test result was 7.1% remain essentially unchanged if "marking A1c as out of range" was removed from the regression equations.

The magnitude of both of these effects is illustrated via the predicted rates of intentions to call one's doctor shown in Figure 3. What is clear from this figure is that health literacy and numeracy directly impacted sensitivity to the test results among participants with diabetes. Our analyses predicted that patients with diabetes with low numeracy skills and low health literacy would be just as likely to call their doctor when the hemoglobin A_{1c} levels were 7.1% or 8.4%. In contrast, our model predicted that highly numerate and health literate patients with diabetes were far more likely (a 34% difference in rates) to call their doctor when the hemoglobin A_{1c}level was 8.4% than 7.1%.

Figure 3. Predicted probabilities that participants with diabetes would call their doctor by reported hemoglobin A1c level and by lower versus higher literacy and numeracy levels.



Perceived Usefulness of Test Results Displays

The 2 questions that measured perceived usefulness of the test results displays were highly correlated (*r*=.77), and the combined measure was highly reliable (Cronbach alpha=.87). Mean perceived usefulness was only 3.38 (SD 1.18) on a 1-5 scale, indicating a significant number of participants found these tables to be difficult to understand and/or not useful. Both numeracy and health literacy were positively correlated with ratings of perceived usefulness (numeracy: *r*=.32, *P*<.001; health literacy: *r*=.26, *P*<.001).

Discussion

Principal Results

A key reason why many patients want direct access to their medical test results is to verify which of their tests are okay and which are not. Unfortunately, our results suggest that many people find the task of identifying out-of-range values embedded in standard test result tables to be quite difficult. Perhaps more importantly, ability to accomplish this task appears highly related to both numeracy and health literacy skills. Participants with lower numeracy and health literacy skills were less than half as likely as those with higher numeracy/literacy abilities to identify hemoglobin A_{1c} levels as outside the reference range in a larger set of results, even though that test was specifically identified in the scenario as the reason for testing. Correctly identifying hemoglobin A_{1c} levels as out of range was, in turn, the single largest predictor of both perception of blood glucose control and intention to call one's doctor in response to the elevated test results.

Health literacy and numeracy skills also appear to enable patients to know when they do or do not need to act in response to test results even after controlling for the effect of being able to correctly identify hemoglobin A_{1c} levels as out of range. Among study participants with diabetes (971/1812, 53.59%), increased health literacy was associated with lower intentions to call the doctor's office for the (barely elevated) test result of hemoglobin A_{1c}=7.1%. When the test result was the more substantially elevated hemoglobin A_{1c} value of 8.4%, it was the more numerate participants who were significantly more likely to call their doctor in response. In contrast, less numerate participants did not appear to recognize that the substantial jump in their hemoglobin A_{1c} results represented a trend worthy of immediate response.

This latter effect is particularly important, as it demonstrates the important distinction between patients knowing their test result numbers versus grasping the personal meaning of those data. For example, a patient with diabetes may use a patient portal to learn that her hemoglobin A_{1c} level changed from 10.1% to 9.3%, but have no idea that a change of less than 1 percentage point represents a significant reduction that corresponds to substantial health and risk reduction benefits. For this patient, knowing the numerical value of her test results did not ensure that she understood what those numbers implied or what actions she needed to consider. Her data were literally meaningless, and she is likely to ignore them in managing her health.

We deliberately studied reactions to laboratory test results among participants with and without diabetes, who would be expected to be more and less familiar with the types of data, respectively. Our results suggest that although familiarity with a metric such as hemoglobin A_{1c} levels is an important first hurdle for new patients, education about these measures is unlikely to be enough to achieve understanding for all patients. Those with low health literacy and low numeracy may require additional support, and interface design for laboratory results in patient portals should take these factors into account.

Relationship to Prior Work

The generally welcome trend in recent years of people gaining access to their own test results has given rise to a common concern about the design of test results displays. In 2010, Wired Magazine ran a feature article titled "The Blood Test Gets a

Makeover," in which several designers were asked to develop "proof of concept" graphical test results reports [43]. These concept graphics used simple line graphs with clear, strong, color cues, reference points, and explanatory language to make multiple types of test results more meaningful to patients. In 2012, the US General Services Administration sponsored a test result design challenge that attracted over 230 entries [44]. These design initiatives demonstrate the widespread concern about this issue and show promise of improvement.

Unfortunately, the design concepts generated have yet to be studied rigorously to evaluate their effects on patient comprehension and activation. For example, do the high/low "flags" often included in clinician interfaces for EHRs (but conspicuously omitted in the patient format used by the major academic medical system we modeled our stimuli after) provide net benefit by clarifying out-of-range values, or do they cause net harm by increasing patient alarm about values that are not clinically concerning? Would use of a categorization system (perhaps with icons) that labeled results by potential harm, not just what is inside or outside the standard reference range values, be useful in guiding patient behavior? Would horizontal line displays or color coding help the less numerate or less literate patients be better able to derive meaning from their test results?

Recent research on risk communication suggests that well-designed visual displays can improve patient understanding of medical data, especially among those with lower numeracy skills [45-47]. Research also supports the supposition that including relevant reference standards beyond the "standard range" values in such displays is likely to make even unfamiliar test data more intuitively interpretable for patients [46,48,49]. Empirical research is needed to answer the preceding questions, thereby guiding the design of results displays to ensure that data are meaningful across levels of literacy and numeracy. In its absence, these barriers will continue to impede effective patient use of test result data (in electronic health records or elsewhere) to improve patient self-management and patient-provider communications.

Yet, it is worth asking ourselves: why are we giving patients these numbers? In many circumstances, patients' informational goals would be addressed more directly by communications that highlight evaluative categories (eg, "poor," "very high," "borderline high") over the specific numerical values. Both we and others have recently argued that precise numerical communications of health data can sometimes be counterproductive [50,51]. Our results suggest that patients with limited numeracy and health literacy skills may be particularly likely to benefit from alternate communication approaches that reinforce the critical "gist" messages [52] before presenting quantitative test result data.

Limitations

Our findings are tempered by several important limitations. First, our study involved a hypothetical vignette and mock test results presented to people who knew they were taking a survey. The lack of personal relevance of these data may have inhibited participants' motivation to seek out and identify the out-of-range values, and it is certainly possible that both perceptions of blood glucose control and intentions to call one's doctor might be different if these were the patient's own test results viewed in an actual EHR portal. Another limitation is that the study displayed all test results simultaneously on a single page (to facilitate their presentation within the survey engine and to allow us to test understanding, not recall), whereas many electronic health record systems only show 1 panel's worth of results at a time. Although both of these limitations may affect the generalizability of our findings into actual clinical practice, this controlled experiment demonstrates the plausibility of literacy and numeracy concerns. Nonetheless, further research is clearly needed to study how well patients understand their own test results in a patient portal.

Conclusions

Our results reinforce the critical role of health literacy and numeracy skills in enabling patients to take active roles in their health care. Being an "informed" patient requires more than having access to test results or being able to recite specific numbers. It means understanding what test data mean for evaluating one's health status and how it should influence future health decisions or behaviors. Our data demonstrate that limited health literacy and numeracy are significant barriers to such knowledge translation tasks. Further research should investigate designs that help people better interpret the meaning of their numbers.

Acknowledgments

The authors would like to thank Jeffrey Kullgren, MD, for assistance in creating the medical scenario. This research was supported by a grant from the University of Michigan Risk Science Center, which is supported by individual donations (including a major gift from the Gelman Educational Foundation) and from project-specific research grants. The funding agreement ensured the authors' independence in designing the study, interpreting the data, and publishing the report.

Authors' Contributions

BJZ-F designed the study, obtained the funding, analyzed the data, and wrote the manuscript. NLE provided feedback on the study design, managed the data collection, and provided critical review of manuscript drafts. HOW provided feedback on study design, data analyses, and manuscript drafts.

Conflicts of Interest

None declared.

References

1. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):S162-S172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]
2. Cimino JJ, Patel VL, Kushniruk AW. The patient clinical information system (PatCIS): technical solutions for and experience with giving patients access to their electronic medical records. Int J Med Inform 2002 Dec 18;68(1-3):113-127. [Medline: 12467796]
3. Keselman A, Slaughter L, Smith CA, Kim H, Divita G, Browne A, et al. Towards consumer-friendly PHRs: patients' experience with reviewing their health records. AMIA Annu Symp Proc 2007:399-403 [FREE Full text] [Medline: 18693866]
4. Bhavnani V, Fisher B, Winfield M, Seed P. How patients use access to their electronic GP record--a quantitative study. Fam Pract 2011 Apr;28(2):188-194 [FREE Full text] [doi: 10.1093/fampra/cmq092] [Medline: 21084568]
5. Laine C, Davidoff F. Patient-centered medicine. A professional evolution. JAMA 1996 Jan 10;275(2):152-156. [Medline: 8531314]
6. Bernat JL, Peterson LM. Patient-centered informed consent in surgical practice. Arch Surg 2006 Jan;141(1):86-92. [doi: 10.1001/archsurg.141.1.86] [Medline: 16415417]
7. Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. Health Aff (Millwood) 2013 Feb;32(2):223-231. [doi: 10.1377/hlthaff.2012.1133] [Medline: 23381514]
8. Wynia M, Dunn K. Dreams and nightmares: practical and ethical issues for patients and physicians using personal health records. J Law Med Ethics 2010;38(1):64-73. [doi: 10.1111/j.1748-720X.2010.00467.x] [Medline: 20446985]
9. Swan M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. Int J Environ Res Public Health 2009 Feb;6(2):492-525 [FREE Full text] [doi: 10.3390/ijerph6020492] [Medline: 19440396]
10. Gilhooly ML, McGhee SM. Medical records: practicalities and principles of patient possession. J Med Ethics 1991 Sep;17(3):138-143 [FREE Full text] [Medline: 1941955]
11. Baldwin DM, Quintela J, Duclos C, Staton EW, Pace WD. Patient preferences for notification of normal laboratory test results: a report from the ASIPS Collaborative. BMC Fam Pract 2005 Mar 8;6(1):11 [FREE Full text] [doi: 10.1186/1471-2296-6-11] [Medline: 15755328]
12. Casalino LP, Dunham D, Chin MH, Bielang R, Kistner EO, Karrison TG, et al. Frequency of failure to inform patients of clinically significant outpatient test results. Arch Intern Med 2009 Jun 22;169(12):1123-1129. [doi: 10.1001/archinternmed.2009.130] [Medline: 19546413]
13. Bonander J, Gates S. Public health in an era of personal health records: opportunities for innovation and new partnerships. J Med Internet Res 2010;12(3):e33 [FREE Full text] [doi: 10.2196/jmir.1346] [Medline: 20699216]
14. McCray AT. Promoting health literacy. J Am Med Inform Assoc 2005;12(2):152-163 [FREE Full text] [doi: 10.1197/jamia.M1687] [Medline: 15561782]
15. 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]
16. Baker DW. The meaning and the measure of health literacy. J Gen Intern Med 2006 Aug;21(8):878-883 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00540.x] [Medline: 16881951]
17. Marks JR, Schectman JM, Groninger H, Plews-Ogan ML. The association of health literacy and socio-demographic factors with medication knowledge. Patient Educ Couns 2010 Mar;78(3):372-376. [doi: 10.1016/j.pec.2009.06.017] [Medline: 19773144]
18. Davis TC, Wolf MS, Bass PF, Thompson JA, Tilson HH, Neuberger M, et al. Literacy and misunderstanding prescription drug labels. Ann Intern Med 2006 Dec 19;145(12):887-894. [Medline: 17135578]
19. Lindquist LA, Go L, Fleisher J, Jain N, Friesema E, Baker DW. Relationship of health literacy to intentional and unintentional non-adherence of hospital discharge medications. J Gen Intern Med 2012 Feb;27(2):173-178 [FREE Full text] [doi: 10.1007/s11606-011-1886-3] [Medline: 21971600]
20. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. Arch Intern Med 2007 Jul 23;167(14):1503-1509. [doi: 10.1001/archinte.167.14.1503] [Medline: 17646604]
21. Wolf MS, Feinglass J, Thompson J, Baker DW. In search of 'low health literacy': threshold vs. gradient effect of literacy on health status and mortality. Soc Sci Med 2010 May;70(9):1335-1341. [doi: 10.1016/j.socscimed.2009.12.013] [Medline: 20167411]
22. Sarkar U, Karter AJ, Liu JY, 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]
23. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]

24. Lipkus IM, Peters E. Understanding the role of numeracy in health: proposed theoretical framework and practical insights. Health Educ Behav 2009 Dec;36(6):1065-1081 [FREE Full text] [doi: 10.1177/1090198109341533] [Medline: 19834054]
25. Nelson W, Reyna VF, Fagerlin A, Lipkus I, Peters E. Clinical implications of numeracy: theory and practice. Ann Behav Med 2008 Jun;35(3):261-274 [FREE Full text] [doi: 10.1007/s12160-008-9037-8] [Medline: 18677452]
26. Kumar D, Sanders L, Perrin EM, Lokker N, Patterson B, Gunn V, et al. Parental understanding of infant health information: health literacy, numeracy, and the Parental Health Literacy Activities Test (PHLAT). Acad Pediatr 2010;10(5):309-316 [FREE Full text] [doi: 10.1016/j.acap.2010.06.007] [Medline: 20674532]
27. Rothman RL, Housam R, Weiss H, Davis D, Gregory R, Gebretsadik T, et al. Patient understanding of food labels: the role of literacy and numeracy. Am J Prev Med 2006 Nov;31(5):391-398. [doi: 10.1016/j.amepre.2006.07.025] [Medline: 17046410]
28. Estrada CA, Martin-Hryniewicz M, Peek BT, Collins C, Byrd JC. Literacy and numeracy skills and anticoagulation control. Am J Med Sci 2004 Aug;328(2):88-93. [Medline: 15311167]
29. Marden S, Thomas PW, Sheppard ZA, Knott J, Lueddeke J, Kerr D. Poor numeracy skills are associated with glycaemic control in Type 1 diabetes. Diabet Med 2012 May;29(5):662-669. [doi: 10.1111/j.1464-5491.2011.03466.x] [Medline: 21978203]
30. Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. Numeracy and decision making. Psychol Sci 2006 May;17(5):407-413. [doi: 10.1111/j.1467-9280.2006.01720.x] [Medline: 16683928]
31. Loewenstein GF, Weber EU, Hsee CK, Welch N. Risk as feelings. Psychol Bull 2001 Mar;127(2):267-286. [Medline: 11316014]
32. Peters E, Lipkus I, Diefenbach MA. The functions of affect in health communications and in the construction of health preferences. J Commun 2006 Aug;56(s1):S140-S162. [doi: 10.1111/j.1460-2466.2006.00287.x]
33. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Risky feelings: why a 6% risk of cancer does not always feel like 6%. Patient Educ Couns 2010 Dec;81 Suppl:S87-S93 [FREE Full text] [doi: 10.1016/j.pec.2010.07.041] [Medline: 20739135]
34. Power TE, Swartzman LC, Robinson JW. Cognitive-emotional decision making (CEDM): a framework of patient medical decision making. Patient Educ Couns 2011 May;83(2):163-169. [doi: 10.1016/j.pec.2010.05.021] [Medline: 20573468]
35. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. J Am Med Inform Assoc 2013 May 1;20(3):526-534 [FREE Full text] [doi: 10.1136/amiajnl-2012-001263] [Medline: 23171659]
36. Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Med Care Res Rev 2007 Apr;64(2):169-190. [doi: 10.1177/10775587070640020301] [Medline: 17406019]
37. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making 2001 Feb;21(1):37-44. [Medline: 11206945]
38. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Aff (Millwood) 2007;26(3):741-748 [FREE Full text] [doi: 10.1377/hlthaff.26.3.741] [Medline: 17485752]
39. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007;27(5):672-680. [doi: 10.1177/0272989X07304449] [Medline: 17641137]
40. Zikmund-Fisher BJ, Angott AM, Ubel PA. The benefits of discussing adjuvant therapies one at a time instead of all at once. Breast Cancer Res Treat 2011 Aug;129(1):79-87 [FREE Full text] [doi: 10.1007/s10549-010-1193-4] [Medline: 20945090]
41. Zikmund-Fisher BJ, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making 2007;27(5):663-671. [doi: 10.1177/0272989X07303824] [Medline: 17652180]
42. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008 May;23(5):561-566 [FREE Full text] [doi: 10.1007/s11606-008-0520-5] [Medline: 18335281]
43. Leckart S. The Wired Magazine. 2010 Nov 29. The blood test gets a makeover URL: http://www.wired.com/2010/11/ff_bloodwork/ [WebCite Cache ID 6RfZp9n6g]
44. d+collab. 2013. The patient record URL: http://healthdesignchallenge.com/ [accessed 2014-07-25] [WebCite Cache ID 6RL9G9qyU]
45. Garcia-Retamero R, Galesic M. Who profits from visual aids: overcoming challenges in people's understanding of risks [corrected]. Soc Sci Med 2010 Apr;70(7):1019-1025. [doi: 10.1016/j.socscimed.2009.11.031] [Medline: 20116159]
46. Keller C, Siegrist M, Visschers V. Effect of risk ladder format on risk perception in high- and low-numerate individuals. Risk Anal 2009 Sep;29(9):1255-1264. [doi: 10.1111/j.1539-6924.2009.01261.x] [Medline: 19572963]
47. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. Patient Educ Couns 2008 Dec;73(3):448-455. [doi: 10.1016/j.pec.2008.07.023] [Medline: 18755566]

48. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. "Is 28% good or bad?" Evaluability and preference reversals in health care decisions. Med Decis Making 2004;24(2):142-148. [doi: 10.1177/0272989X04263154] [Medline: 15090100]
49. Windschitl PD, Martin R, Flugstad AR. Context and the interpretation of likelihood information: the role of intergroup comparisons on perceived vulnerability. J Pers Soc Psychol 2002 May;82(5):742-755. [Medline: 12003474]
50. Zikmund-Fisher BJ. The right tool is what they need, not what we have: a taxonomy of appropriate levels of precision in patient risk communication. Med Care Res Rev 2013 Feb;70(1 Suppl):37S-49S. [doi: 10.1177/1077558712458541] [Medline: 22955699]
51. Schwartz PH. Questioning the quantitative imperative: decision aids, prevention, and the ethics of disclosure. Hastings Cent Rep 2011;41(2):30-39. [Medline: 21495515]
52. Reyna VF. How people make decisions that involve risk: a dual-processes approach. Curr Dir Psychol Sci 2004 Apr;13(2):60-66. [doi: 10.1111/j.0963-7214.2004.00275.x]

Abbreviations

CBC: complete blood cell count
EHR: electronic health record
MCH: mean corpuscular hemoglobin
MCHC: mean corpuscular hemoglobin concentration
REALM: Rapid Estimate of Adult Literacy in Medicine
S-TOFHLA: Short Test of Functional Health Literacy in Adults
SNS: Subjective Numeracy Scale
SSI: Survey Sampling International
WBC: white blood cell

Edited by G Eysenbach; submitted 18.01.14; peer-reviewed by R Britt, D Schillinger; comments to author 26.04.14; revised version received 07.06.14; accepted 10.07.14; published 08.08.14.

Please cite as:
Zikmund-Fisher BJ, Exe NL, Witteman HO
Numeracy and Literacy Independently Predict Patients' Ability to Identify Out-of-Range Test Results
J Med Internet Res 2014;16(8):e187
URL: http://www.jmir.org/2014/8/e187/
doi:10.2196/jmir.3241
PMID:25135688

Original Paper

Online Dietary Intake Estimation: Reproducibility and Validity of the Food4Me Food Frequency Questionnaire Against a 4-Day Weighed Food Record

Rosalind Fallaize[1], BSc (Hons); Hannah Forster[2], BSc (Hons); Anna L Macready[1], BSc (Hons), MSc, PhD; Marianne C Walsh[2], BSc (Hons), PhD; John C Mathers[3], BSc (Hons), PhD; Lorraine Brennan[2], BA (Mod), PhD; Eileen R Gibney[2], BSc (Hons), MSc, PhD; Michael J Gibney[2], BAgrSc, MAgrSc, PhD; Julie A Lovegrove[1], BSc (Hons), PhD

[1]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, Department of Food and Nutritional Sciences, University of Reading, Reading, United Kingdom

[2]UCD Institute of Food and Health, UCD Centre for Molecular Innovation, University College Dublin, Dublin, Ireland

[3]Human Nutrition Research Centre, Institute for Health and Ageing, Newcastle University, Newcastle, United Kingdom

Corresponding Author:
Julie A Lovegrove, BSc (Hons), PhD
Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research
Department of Food and Nutritional Sciences
University of Reading
Whiteknights
PO Box 226
Reading, RG6 6AP
United Kingdom
Phone: 44 118 378 6418
Fax: 44 118 378 7708
Email: j.a.lovegrove@reading.ac.uk

Abstract

Background: Advances in nutritional assessment are continuing to embrace developments in computer technology. The online Food4Me food frequency questionnaire (FFQ) was created as an electronic system for the collection of nutrient intake data. To ensure its accuracy in assessing both nutrient and food group intake, further validation against data obtained using a reliable, but independent, instrument and assessment of its reproducibility are required.

Objective: The aim was to assess the reproducibility and validity of the Food4Me FFQ against a 4-day weighed food record (WFR).

Methods: Reproducibility of the Food4Me FFQ was assessed using test-retest methodology by asking participants to complete the FFQ on 2 occasions 4 weeks apart. To assess the validity of the Food4Me FFQ against the 4-day WFR, half the participants were also asked to complete a 4-day WFR 1 week after the first administration of the Food4Me FFQ. Level of agreement between nutrient and food group intakes estimated by the repeated Food4Me FFQ and the Food4Me FFQ and 4-day WFR were evaluated using Bland-Altman methodology and classification into quartiles of daily intake. Crude unadjusted correlation coefficients were also calculated for nutrient and food group intakes.

Results: In total, 100 people participated in the assessment of reproducibility (mean age 32, SD 12 years), and 49 of these (mean age 27, SD 8 years) also took part in the assessment of validity. Crude unadjusted correlations for repeated Food4Me FFQ ranged from .65 (vitamin D) to .90 (alcohol). The mean cross-classification into “exact agreement plus adjacent” was 92% for both nutrient and food group intakes, and Bland-Altman plots showed good agreement for energy-adjusted macronutrient intakes. Agreement between the Food4Me FFQ and 4-day WFR varied, with crude unadjusted correlations ranging from .23 (vitamin D) to .65 (protein, % total energy) for nutrient intakes and .11 (soups, sauces and miscellaneous foods) to .73 (yogurts) for food group intake. The mean cross-classification into “exact agreement plus adjacent” was 80% and 78% for nutrient and food group intake, respectively. There were no significant differences between energy intakes estimated using the Food4Me FFQ and 4-day WFR, and Bland-Altman plots showed good agreement for both energy and energy-controlled nutrient intakes.

Conclusions: The results demonstrate that the online Food4Me FFQ is reproducible for assessing nutrient and food group intake and has moderate agreement with the 4-day WFR for assessing energy and energy-adjusted nutrient intakes. The Food4Me FFQ is a suitable online tool for assessing dietary intake in healthy adults.

(J Med Internet Res 2014;16(8):e190) doi:10.2196/jmir.3355

KEYWORDS

food frequency questionnaire; weighed food record; validity; reproducibility; dietary assessment; Food4Me; Web-based

Introduction

Given the continuing rise in some noncommunicable diseases and the growing burden of diet-related ill health [1-3], researchers are seeking new and innovative ways of facilitating dietary change. These include the application of digital technologies, which are revolutionizing the delivery of health-related services because of their reduced costs and wide reach. Online interventions are particularly promising because they have the potential to increase exposure to health promotion material. Recent estimates show that Internet use has increased by >150% in North America and by nearly 400% in Europe since 2000, with a total of 78.6% and 63.2% of these populations, respectively, now classified as Internet users [4]. Given their lower costs, Internet-based services have the potential to enhance the cost-benefit ratio for interventions aimed at prevention of diet-related noncommunicable diseases [5-6]. Furthermore, interactive Web-based interventions have been shown to increase patient activation and self-management capabilities in chronically ill adults [7] and enhance weight loss in obese individuals (compared with non-Web-based interventions) [8].

To quantify dietary change in response to an intervention, an accurate and validated means of assessing food intake is essential [9]. Population-level food intake is usually assessed in 1 of 3 ways: a food frequency questionnaire (FFQ), 24-hour recall, or estimated or weighed food record (WFR). The WFR, which involves weighing all foods and drinks consumed over a 3-7 day period, is often considered the most accurate measure of intake and has been referred to as the imperfect gold standard [10]. However, prospective recording of food consumption can alter the type and quantity of foods eaten and, therefore, introduce bias into the estimate of food intake [11-13]. The FFQ and 24-hour recall, which rely on retrospective recording of food consumption, are also prone to reporting bias, including overestimated consumption of “healthy” foods, such as fruit and vegetables, and underestimation of “unhealthy” food intake. WFR require participants to be highly motivated and are labor-intensive for both participants and researchers. Conversely, FFQ are inexpensive to process and can be self-administered electronically, making them suitable for online interventions. Other advantages include reducing paper use, postage costs, and the space; security; and organization required for paper file storage [14]. For this reason, FFQ are most commonly used in large-scale epidemiological and intervention studies to determine food and nutrient intake [15].

The present research was conducted as part of the Food4Me study, which aims to test the utility of online personalized dietary advice using an online FFQ to assess dietary intake [16,17]. The Food4Me FFQ includes 157 food items and food portion photographs and has been described previously by Forster et al [18]. FFQ are generally validated against existing dietary assessment methods, such as WFR [19], and several FFQ have been validated for electronic and online use recently [14,20-22].

The Food4Me FFQ has been shown to have good agreement with the European Prospective Investigation of Cancer (EPIC)-Norfolk FFQ for the estimation of energy-adjusted nutrient intakes [18]. The aim of this study is to further validate the Food4Me FFQ against a WFR and to assess its reproducibility using a test-retest methodology.

Methods

Study Sample

To accurately estimate the Bland-Altman limits of agreement between 2 methods, a sample size of 50-100 is required [23]. Allowing for 20% dropout, 121 participants aged ≥18 years were recruited from the University of Reading, UK, via email and poster advertising. Participants were provided with a study information sheet before participation and were asked to sign an informed consent form. A participant information form, which included self-reported weight and height measurements, was used to assess suitability for the study. Individuals reporting health issues or ill health, self-reported or diagnosed food intolerances, or special nutritional requirements (eg, pregnancy or lactation) were ineligible to participate. Ethical approval for the study was obtained from the School of Chemistry, Food and Pharmacy Research Ethics Committee, University of Reading, UK (01-12-Lovegrove).

Study Design

Reproducibility of the Food4Me FFQ was determined by asking participants to complete the questionnaire on 2 occasions 4 weeks apart, mimicking its application in the Food4Me study. To assess the validity of the FFQ against a 4-day WFR, half the sample (those recruited first) were asked to complete a 4-day WFR 1 week following the first administration of the Food4Me-FFQ. Participants who completed both the Food4Me FFQ and 4-day WFR were also asked to complete a dietary record usability-rating questionnaire on Survey Monkey (Survey Monkey Inc, Palo Alto, CA, USA) in the week following the completion of the second Food4Me FFQ. The usability-rating questionnaire included questions about ease of use and willingness to complete the records. Participants were asked not to change their diet during the study.

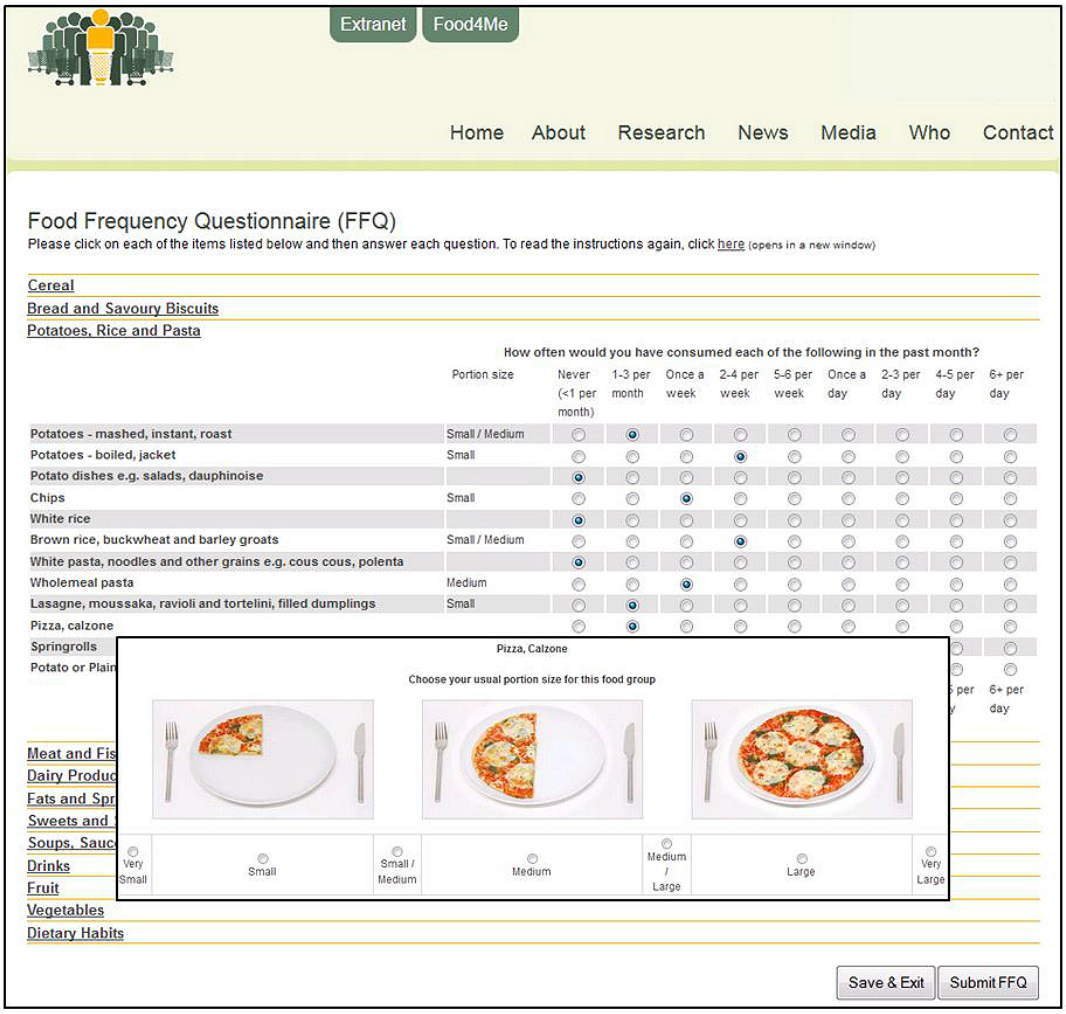
Weighed Food Record

Participants were asked to record all foods and beverages consumed over a nonconsecutive 4-day period that included 3 weekdays (Monday to Thursday) and 1 weekend day (Saturday to Sunday). Before completing the WFR, participants were coached on how to describe food products by a dietitian and provided with weighing scales (Salter Disc Electronic Kitchen Scales SKU# 1036 WHSSDR). When participants were unable to provide weighed portion size information, this was estimated retrospectively within 1 week using the Ministry of Agriculture, Fisheries and Food Portion Size Atlas [24].

Food4Me FFQ

The self-administered Food4Me FFQ is an online, semiquantitative food frequency questionnaire (developed by University College Dublin and Crème Software Ltd). To complete the questionnaire, participants were provided with a website address and unique log-in details. On logging into the server, participants were directed to a webpage containing detailed instructions on how to complete the FFQ. The questionnaire contained questions on the average consumption of 157 food items over the previous month. The food items were divided into the following 11 categories: (1) cereal, (2) bread and savory biscuits, (3) potatoes, rice and pasta, (4) meat and fish, (5) dairy products and fat, (6) fats and spreads, (7) sweets and snacks, (8) soups, sauces and spreads, (9) drinks, (10) fruit, and (11) vegetables. During completion of the Food4Me FFQ, participants were required to provide information on frequency of consumption and portion size. Frequency of consumption was measured by selecting one of the following options: never or less than once a month, 1-3 times a month, once a week, 2-4 times a week, 5-6 times per week, once a day, 2-3 times per day, 5-6 times per day, and >6 times per day. Food portion size was estimated using photographs. Each food item had 3 photographs representing small, medium, and large portions and these descriptors were provided below the appropriate image. Participants could select one of the following options: very small, small, small/medium, medium, medium/large, large, or very large which were linked electronically to portion sizes (in grams) (see Figure 1). Food intake (g/day) was calculated by multiplying frequency of consumption by the specified portion size (see Forster et al for detailed methods [18]). Further screenshots of the online Food4Me FFQ are shown in Multimedia Appendix 1.

Figure 1. Screenshot of the online Food4Me food frequency questionnaire.



Food Frequency Questionnaire (FFQ)

Please click on each of the items listed below and then answer each question. To read the instructions again, click here (opens in a new window)

Cereal

Bread and Savoury Biscuits

Potatoes, Rice and Pasta

How often would you have consumed each of the following in the past month?

	Portion size	Never (<1 per month)	1-3 per month	Once a week	2-4 per week	5-6 per week	Once a day	2-3 per day	4-5 per day	6+ per day
Potatoes - mashed, instant, roast	Small / Medium	○	●	○	○	○	○	○	○	○
Potatoes - boiled, jacket	Small	○	○	○	●	○	○	○	○	○
Potato dishes e.g. salads, dauphinoise		●	○	○	○	○	○	○	○	○
Chips	Small	○	○	●	○	○	○	○	○	○
White rice		●	○	○	○	○	○	○	○	○
Brown rice, buckwheat and barley groats	Small / Medium	○	○	○	●	○	○	○	○	○
White pasta, noodles and other grains e.g. cous cous, polenta		●	○	○	○	○	○	○	○	○
Wholemeal pasta	Medium	○	○	●	○	○	○	○	○	○
Lasagne, moussaka, ravioli and tortelini, filled dumplings	Small	○	●	○	○	○	○	○	○	○
Pizza, calzone		○	●	○	○	○	○	○	○	○
Springrolls									○	○
Potato or Plain									○	○

Pizza, Calzone

Choose your usual portion size for this food group

○ Very Small ○ Small ○ Small / Medium ○ Medium ○ Medium / Large ○ Large ○ Very Large

Meat and Fis

Dairy Produc

Fats and Spr

Sweets and

Soups, Sauc

Drinks

Fruit

Vegetables

Dietary Habits

Dietary Record Usability-Rating Questionnaire

The dietary record usability-rating questionnaire was comprised of 5 questions about the completion of the Food4Me FFQ and the 4-day WFR. Participants who completed both the Food4Me FFQ and the 4-day WFR (n=49) were asked to select one of the following responses to indicate their level of agreement: strongly agree, agree, neither agree nor disagree, disagree, or strongly disagree to the following questions:

1. I found the Food4Me FFQ / 4-day WFR easy to complete
2. I found the Food4Me FFQ / 4-day WFR too time consuming
3. I found the Food4Me FFQ / 4-day WFR interesting to complete
4. I found the Food4Me FFQ / 4-day WFR made me reflect on my intake
5. In the future I would be willing to complete more Food4Me FFQ/ 4-day WFR

Misreporting

The Henry equation [25] was used to calculate basal metabolic rate (BMR), and BMR was multiplied by a physical activity level (PAL) of 1.1 to calculate the lowest possible estimated energy requirements (EER) for each participant [26]. Participants reporting energy intakes lower than their EER were classified as underreporters. Participants reporting a daily energy intake greater than 4500 kcal, which is considered implausibly high, were excluded from the analysis [27].

Nutritional Intake Analysis

Estimated nutritional intake data from the Food4Me FFQ were generated automatically by the online Food4Me programmed system, as described by Forster et al [18]. Composition of the food items listed in the FFQ were derived from WISP (Tinuviel Software, Anglesey, UK) [28] and modified to include recipes of composite dishes, generic commercial foods, new foods on the market, and current manufacturers information. The 4-day WFR intakes were analyzed using WISP (Tinuviel Software, Anglesey, UK) [28]. For the purpose of the current study, consumption of dietary supplements was not included in the analyses.

Statistical Analysis

Statistical analyses were performed using SPSS version 20.0 (IBM Corp, Armonk, NY, USA). Mean nutrient intakes and standard deviations were calculated for baseline characteristics, repeated Food4Me FFQ, and 4-day WFR. Differences in participant characteristics and energy intakes (kcal) were assessed using a paired 2-sample *t* test. Nutrient intakes were compared using general linear model (GLM) analysis controlling for energy and gender where there was significant interaction between gender and nutrient intake. Data were checked for normality using the Shapiro-Wilk test and, depending on the outcome, the association between dietary intake methods and repeated Food4Me FFQ were assessed using either Pearson product-moment correlation or Spearman correlation coefficient (SCC, rho). A *P* value of <.05 was considered statistically significant.

The relative agreement between the dietary intake methods and repeated Food4Me FFQ was assessed using cross-classification of nutrient intakes to estimate the percentage of participants classified into quartiles as follows: exact agreement (percentage of cases classified into the same quartile), exact agreement plus adjacent (percentage of cases cross-classified into the same or adjacent quartile), disagreement (percentage of cases cross-classified 2 quartiles apart), and extreme disagreement (percentage of cases cross-classified into extreme quartiles). For intakes of energy and macronutrients, the Bland-Altman method [29] was used to further assess the limits of agreement between the 2 methods (Food4Me FFQ and WFR) and between the repeated Food4Me FFQ. As per the Bland-Altman methodology, dietary records were considered comparable/ repeatable if greater than 95% of data plots lay within 2 standard deviations of the mean. GraphPad PRISM version 6 was used to produce the Bland-Altman plots (GraphPad Software, Inc, La Jolla, CA, USA).

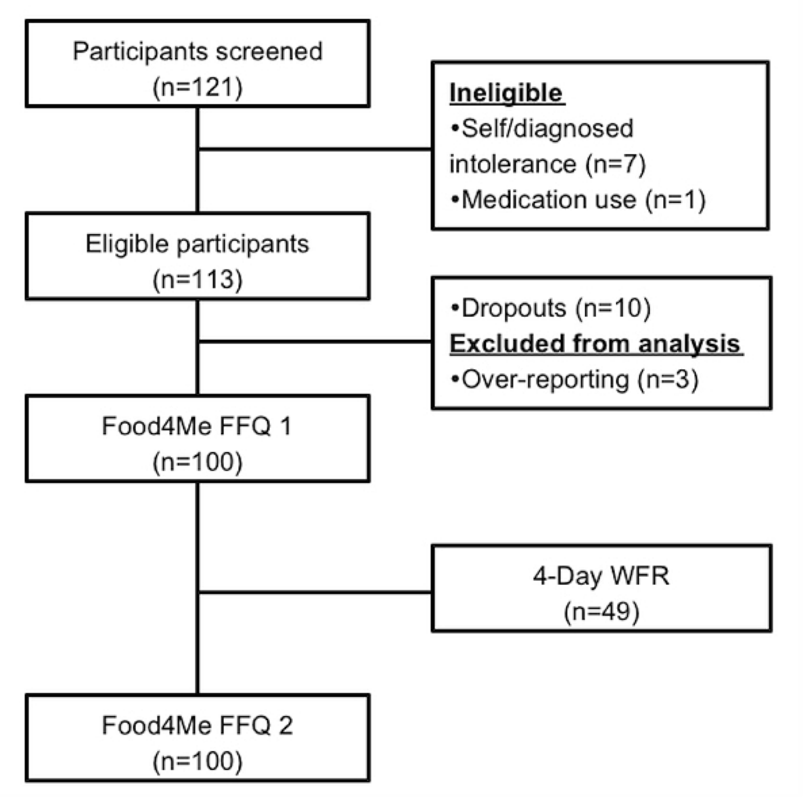
Differences in food group intakes between the FFQ and WFR and repeated Food4Me FFQ were also examined. For this purpose, food items in the Food4Me FFQ and 4-day WFR were arranged into 35 food groups as per previous validation by Forster et al [18]. SCC were calculated to assess the strength of association between methods for estimated intakes of the 35 food groups. To assess the relative agreement between the dietary methods and repeated FFQ for daily food group intake, food groups were also cross-classified to estimate the percentage of participants classified by the 2 methods into quartiles of exact agreement, exact agreement plus adjacent, disagreement, and extreme disagreement.

Results

Summary

A total of 121 participants were screened for inclusion in the study, of which 113 were deemed eligible. Reasons for exclusion included self-diagnosed food intolerance (n=7) and medication use (n=1). Before completion, 10 participants dropped out of the study and a further 3 were excluded from analysis due to a reported energy intake >4500 kcal [26]. The final dataset for analysis included 100 participants, of whom 49 had also completed the 4-day WFR, as illustrated in Figure 2. A total of 48 participants completed the Diet Record Usability-Rating Questionnaire.

Figure 2. Flow of participants through the study.



Overview of the Study Population

Self-reported demographic characteristics of the participants in the reproducibility and validation study are shown in Table 1.

No significant differences were observed between age and body mass index (BMI) for males and females. Participants who completed the WFR (validation study) were, on average, 4.6 years younger than the participant group as a whole.

Table 1. Demographic characteristics of the participants who completed the validation and reproducibility studies according to gender.

Study	N	Demographic characteristics, mean (SD)	
		Age (years)	BMI (kg/m^2)[a]
Reproducibility			
Males	31	30.1 (12.7)	24.3 (3.1)
Females	69	32.1 (11.6)	23.3 (3.3)
All	100	31.5 (11.9)	23.6 (3.3)
Validation			
Males	15	24.2 (7.6)	23.1 (3.2)
Females	34	27.9 (8.6)	22.2 (2.6)
All	49	26.9 (8.4)	22.5 (2.8)

[a] BMI based on self-reported weight and height.

Reproducibility of the Food4Me FFQ

Comparison of Nutrient Intakes Between Repeated Food4Me FFQ

Mean energy and nutrient intakes estimated by repeated measures of the Food4Me FFQ (FFQ1 and FFQ2) are presented in Table 2. Estimated energy intakes were significantly higher in the first administration of the FFQ compared with the second administration (difference=135 kcal/day, equivalent to 6.5% higher, $P<.05$). With the exception of carbohydrate, no significant differences were observed between macronutrient and micronutrient intakes estimated by FFQ1 and FFQ2. Overall, the Food4Me FFQ showed good reproducibility for energy-adjusted nutrient intakes. A total of 16 participants were found to underreport in both FFQ with a further 3 underreporting in FFQ1 and 5 in FFQ2. The removal of underreporters from both FFQ (n=24) did not impact on the reproducibility of the questionnaire (data not shown).

Bland-Altman plots for estimates of energy (kcal), total fat (% total energy, TE), protein (%TE) and carbohydrate (%TE) intakes are shown in Figure 3. The Food4Me FFQ showed good reproducibility for the estimation of daily protein intake, with less than 5% of cases falling outside of the limits of agreement. For energy and total fat intake, 6% of cases fell outside of the limits of agreement and for carbohydrate 7%, indicating similar reproducibility. The mean difference (bias) between energy intakes was relatively small (135 kcal/day) with greater values being estimated at FFQ1. Conversely, estimates of energy-adjusted protein and total fat intake were higher at FFQ2 with biases of –0.22 %TE and –1.23 %TE, respectively. In contrast with the energy-adjusted macronutrient intakes, variation between estimates of energy increased with higher mean energy intakes (Figure 3), suggesting poorer reproducibility for those participants reporting higher energy intakes.

Correlation coefficients for estimates of energy and nutrient intakes between repeated administrations of the Food4Me FFQ are shown in Table 3. Correlation coefficients ranged from .65 (vitamin D) to .90 (alcohol) with a mean value of .75. Correlations were significant for all nutrients (*P*<.01). The cross-classification of quartiles of mean estimated daily energy and nutrient intakes between repeated administrations of the Food4Me FFQ is also shown in Table 3. The percentage of participants classified into quartiles of exact agreement ranged from 45% (polyunsaturated fatty acids %TE) to 74% (vitamin A retinol equivalents, RE). For classifications of exact agreement plus adjacent, values were consistently high, ranging from 87% (vitamin D) to 98% (vitamin A RE). The mean percentage of participants classified into quartiles of disagreement was 7% with less than 1% of participants on average classified into extreme disagreement.

Table 2. Mean daily energy and nutrient intakes estimated by repeated measures of the online Food4Me FFQ and general linear model (GLM) results (N=100).

Nutrient[a]	Questionnaire, mean (SD)		GLM analysis, *P*	
	Food4Me FFQ1	Food4Me FFQ2	Controlled for energy	Controlled for energy and gender[b]
Energy (kcal)	2223.8 (766.2)	2088.8 (705.4)	.008[c]	—
Total fat (g)	85.3 (33.8)	82.0 (32.5)	.24	.24
Total fat (%TE)	34.1 (4.9)	35.1 (6.1)	.21	.21
SFA (g)	33.2 (14.2)	32.2 (14.5)	.13	.13
SFA (%TE)	13.2 (2.4)	13.6 (3.0)	.23	.23
MUFA (g)	32.0 (13.4)	30.6 (12.7)	.38	.38
MUFA (%TE)	12.8 (2.5)	13.1 (2.8)	.36	.36
PUFA (g)	14.4 (5.6)	13.7 (5.3)	.91	.98
PUFA (%TE)	5.8 (1.27)	6.0 (1.5)	.48	.62
Omega 3 (g)	1.70 (0.71)	1.61 (0.66)	.92	.92
Protein (g)	90.6 (35.2)	84.4 (30.8)	.70	.70
Protein (%TE)	16.3 (2.8)	16.4 (4.5)	.84	.84
Carbohydrate (g)	263.9 (87.6)	238.6 (89.6)	.06	.03
Carbohydrate (%TE)	45.1 (6.5)	43.1 (7.9)	.05	.03
Total sugars (g)	125.6 (49.3)	115.9 (56.1)	.65	.79
Total sugars (%TE)	21.4 (5.8)	20.6 (5.8)	.31	.32
Alcohol (g)	12.9 (16.2)	12.8 (15.7)	.61	.60
Calcium (mg)	1085.0 (378.1)	1008.0 (416.5)	.59	.64
Total folate (µg)	361.0 (120.0)	335.9 (416.5)	.43	.43
Iron (mg)	14.9 (5.3)	13.4 (4.3)	.05	.05
Total carotene (µg)	6209.6 (4590.8)	5482.1 (3645.2)	.43	.60
Riboflavin (mg)	2.24 (0.80)	2.11 (0.89)	.87	.98
Thiamin (mg)	2.87 (2.80)	2.91 (3.33)	.76	.76
Vitamin B6 (mg)	2.54 (0.86)	2.35 (0.78)	.31	.31
Vitamin B12 (µg)	7.32 (3.57)	6.72 (3.60)	.64	.64
Vitamin C (mg)	167.8 (82.5)	153.8 (74.7)	.42	.51
Vitamin A RE (µg)	1160.9 (1015.8)	1057.8 (907.9)	.79	.79
Retinol (µg)	502.9 (408.0)	470.1 (400.5)	.99	.99
Vitamin D (µg)	3.89 (2.39)	3.51 (1.90)	.48	.38
Vitamin E (mg)	10.61 (4.05)	9.77 (3.86)	.39	.29
Salt (g)	6.30 (2.70)	5.92 (2.24)	.97	.97

[a] MUFA: monounsaturated fatty acids; PUFA: polyunsaturated fatty acids; RE: retinol equivalents; SFA: saturated fatty acids; TE: total energy.

[b] Controlled for gender where appropriate. No significant interactions were observed between method and gender.

[c] *P* value derived using 2-samples paired *t* test.

Table 3. Unadjusted correlation coefficients and cross-classification of quartiles of mean energy and nutrient intakes derived from repeat measures of the online Food4Me FFQ (N=100).

Nutrient[a]	Correlation[b]	Quartiles, %			
		Exact agreement[c]	Exact agreement plus adjacent[d]	Disagreement[e]	Extreme disagreement[f]
Energy (kcal)	.77[g]	57	90	9	1
Total fat (g)	.81	64	92	8	0
Total fat (%TE)	.72	56	91	7	2
SFA (g)	.81	60	91	9	0
SFA (%TE)	.70	46	88	12	0
MUFA (g)	.80	51	96	4	0
MUFA (%TE)	.70	53	89	10	1
PUFA (g)	.78	51	95	3	2
PUFA (%TE)	.68	45	92	6	2
Omega 3 (g)	.78	58	91	9	0
Protein (g)	.80	56	88	12	0
Protein (%TE)	.73	59	93	7	0
Carbohydrate (g)	.74	53	96	4	0
Carbohydrate (%TE)	.73	62	89	9	2
Total sugars (g)	.77	66	94	4	2
Total sugars (%TE)	.69	61	88	11	1
Alcohol (g)	.90	70	96	4	0
Calcium (mg)	.73	55	92	7	1
Total folate (µg)	.74	53	93	6	1
Iron (mg)	.75	53	95	4	1
Total carotene (µg)	.76	60	90	10	0
Riboflavin (mg)	.73	56	90	8	2
Thiamin (mg)	.71	51	91	6	3
Vitamin B6 (mg)	.72	56	89	10	1
Vitamin B12 (µg)	.73	64	95	5	0
Vitamin C (mg)	.72	60	95	5	0
Vitamin A (RE) (µg)	.90	74	98	2	0
Retinol (µg)	.67	50	90	7	3
Vitamin D (µg)	.65	52	87	12	1
Vitamin E (mg)	.75	56	91	7	2
Salt (g)	.78	57	90	8	2

[a] MUFA: monounsaturated fatty acids; PUFA: polyunsaturated fatty acids; RE: retinol equivalents; SFA: saturated fatty acids; TE: total energy.

[b] Correlation is significant at the .01 level (2-tailed) for all nutrients analyzed.

[c] Exact agreement, % of cases cross-classified into the same quartile.

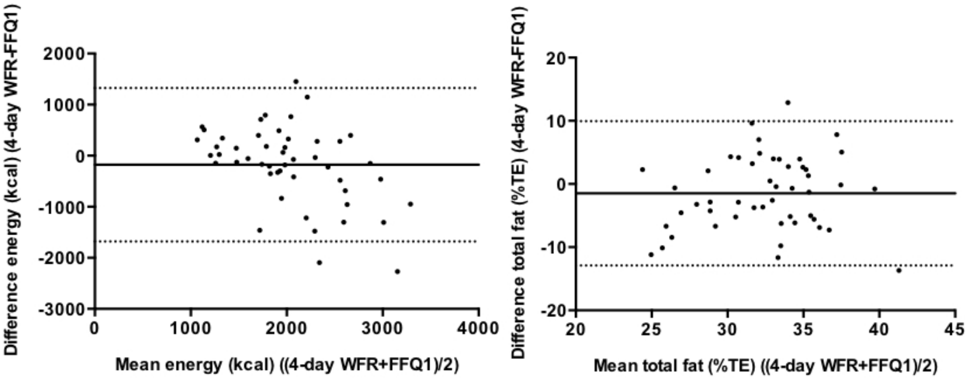
[d] Exact agreement plus adjacent, % of cases cross-classified into the same or adjacent quartile.

[e] Disagreement, % of cases cross-classified 2 quartiles apart.

[f] Extreme disagreement, % of cases cross-classified into extreme quartiles.

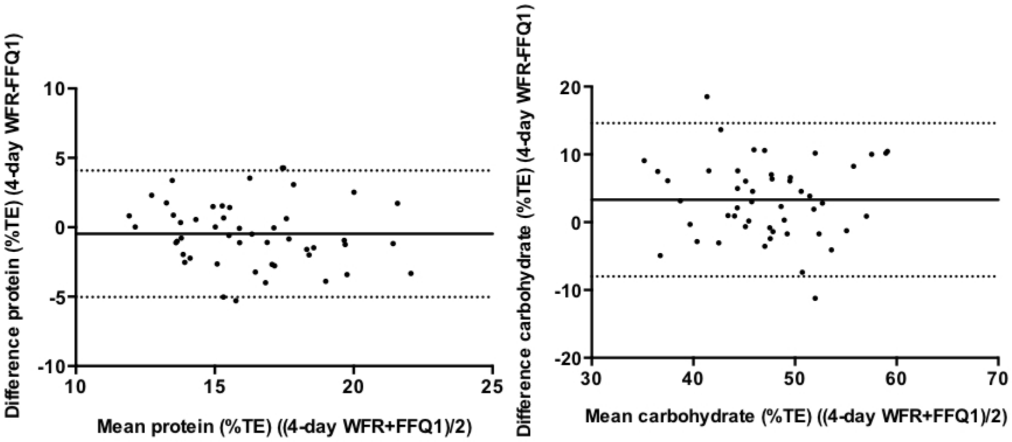
[g] Pearson correlation.

Figure 3. Reproducibility study Bland-Altman plots for (a) energy, (b) total fat, (c) protein, and (d) carbohydrate with the mean difference and limits of agreement. The solid line represents the mean difference and the dotted lines represent the limits of agreement.



a. Energy (kcal)

b. Total Fat (%TE)



c. Protein (%TE)

d. Total Carbohydrate (%TE)

Comparison of Food Group Intakes Between Repeated Food4Me FFQ

To assess differences in food group intake between repeated administrations of the online Food4Me FFQ, food items were categorized into 35 food groups. Correlation coefficients and cross-classification of mean food group intakes are presented in Table 4. SCC ranged from .55 (tinned fruit or vegetables) to .92 (alcoholic beverages) with a mean value of .75. Correlations were significant for all food groups (*P*<.01). The percentage of participants classified into quartiles of exact agreement ranged from 46% (fats and oils) to 86% (tinned fruit or vegetables). For classifications of exact agreement plus adjacent values were consistently high, ranging from 81% (eggs and egg dishes) to 99% (alcoholic beverages). The mean percentage of participants classified into quartiles of disagreement was 7% and for extreme disagreement was 1%.

Table 4. Spearman correlation coefficients (SCC) and cross-classification of quartiles of food group intake derived from repeat measures of the online Food4Me FFQ (n=100).

Food group	SCC[a]	Quartile, %			
		Exact agreement[b]	Exact agreement plus adjacent[c]	Disagreement[d]	Extreme disagreement[e]
Rice, pasta, grains and starches	.78	56	92	8	0
Savories (lasagne, pizza)	.70	52	89	9	2
White bread (rolls, tortillas, crackers)	.83	62	94	6	0
Wholemeal, brown breads, and rolls	.77	60	91	8	1
Breakfast cereals and porridge	.90	67	96	4	0
Biscuits	.56	48	86	11	3
Cakes, pastries and buns	.64	52	87	10	3
Milk	.74	49	91	7	2
Cheeses	.66	52	87	11	2
Yogurts	.79	58	97	2	1
Ice cream, creams and desserts	.77	55	94	4	2
Eggs and egg dishes	.69	56	81	16	3
Fats and oils (eg, butter, low-fat spreads, hard cooking fats)	.64	46	89	8	3
Potatoes and potato dishes	.61	51	86	9	5
Chipped, fried & roasted potatoes	.61	57	87	12	1
Peas, beans and lentils and vegetable and pulse dishes	.75	62	92	7	1
Green vegetables	.76	54	93	6	1
Carrots	.68	54	90	8	2
Salad vegetables (eg, lettuce)	.77	58	92	6	2
Other vegetables (eg, onions)	.85	56	97	2	1
Tinned fruit or vegetables	.55	86	86	14	0
Bananas	.81	60	95	5	0
Other fruits (eg, apples, pears, oranges)	.86	61	97	3	0
Nuts and seeds, herbs and spices	.77	68	84	13	3
Fish and fish products/dishes	.84	57	95	5	0
Bacon and ham	.88	73	97	3	0
Red meat (eg, beef, veal, lamb, pork)	.74	67	90	9	1
Poultry (chicken and turkey)	.75	54	93	6	1
Meat products (eg, burgers, sausages, pies, processed meats)	.85	62	93	7	0
Alcoholic beverages	.92	71	99	1	0
Sugars, syrups, preserves, and sweeteners	.78	80	94	6	0
Confectionary and savory snacks	.73	52	92	8	0
Soups, sauces, and miscellaneous foods	.69	59	90	8	2
Teas and coffees	.85	69	96	3	1
Other beverages (eg, fruit juices, carbonated beverages, squash)	.75	54	95	4	1

[a] Correlation is significant at the .01 level (2-tailed) for all nutrients analyzed.

[b] Exact agreement, % of case cross-classified into the same quartile.

[c] Exact agreement plus adjacent, % of cases cross-classified into the same or adjacent quartile.

[d] Disagreement, % of cases cross-classified 2 quartiles apart.

[e] Extreme disagreement, % of cases cross-classified into extreme quartiles.

Validation of the Food4Me FFQ

Comparison of Nutrient Intakes Between the Food4Me FFQ and 4-Day WFR

Mean energy and nutrient intakes estimated by the 4-day WFR and Food4Me FFQ (FFQ1) are presented in Table 5. There were no significant differences between estimates of energy intake. However, when underreporters had been removed (n=19), a significant difference in intake was observed ($P<.05$) (Table 5). In total, 12 participants underreported in the 4-day WFR and 15 underreported in the Food4Me FFQ1, with 8 of these underreporting in both methods.

Table 5. Mean daily energy and nutrient intakes estimated by online Food4Me FFQ and 4-day WFR and general linear model (GLM) results (n=49).

Nutrient[a]	Questionnaire, mean (SD)		GLM analysis, *P*	
	Food4Me FFQ1	4-day WFR	Controlled for energy	Controlled for energy and gender[b]
Energy (kcal)	2115.2 (809.1)	1936.9 (505.8)	.11[c]	—
Total fat (g)	79.6 (36.2)	68.6 (22.2)	.10	.10
Total fat (%TE)	33.1 (4.5)	31.6 (5.1)	.13	.13
SFA (g)	45.6 (15.6)	24.3 (10.4)	<.001	<.001
SFA (%TE)	13.1 (2.3)	11.0 (2.9)	<.001	<.001
MUFA (g)	29.8 (4.5)	21.4 (7.3)	<.001	<.001
MUFA (%TE)	12.4 (2.6)	9.8 (2.0)	<.001	<.001
PUFA (g)	12.7 (4.9)	10.7 (4.6)	.12	.12
PUFA (%TE)	5.44 (0.9)	4.97 (1.6)	.10	.10
Protein (g)	87.2 (36.0)	77.2 (21.4)	.31	.31
Protein (%TE)	16.5 (2.9)	16.1 (2.6)	.40	.40
Carbohydrate (g)	253.4 (94.1)	248.3 (54.9)	.20	.20
Carbohydrate (%TE)	45.6 (6.6)	48.9 (6.5)	.01	.01
Total sugars (g)	119.1 (46.7)	102.8 (37.8)	.18	.18
Total sugars (%TE)	21.5 (5.5)	20.1 (6.2)	.25	.25
Alcohol (g)	13.0 (14.5)	11.6 (22.2)	.50	.50
Calcium (mg)	1043.8 (386.8)	865.8 (285.5)	.003	.001
Total folate (µg)	337.6 (124.6)	273.8 (139.5)	.05	.11
Iron (mg)	14.1 (5.4)	13.0 (5.6)	.98	.98
Total carotene (µg)	5011.4 (3321.2)	2725.3 (2995.3)	.001	.001
Riboflavin (mg)	2.27 (0.83)	1.85 (0.82)	.04	.04
Thiamin (mg)	2.22 (1.56)	2.19 (3.26)	.98	.98
Vitamin B6 (mg)	2.44 (0.83)	2.09 (0.70)	.06	.06
Vitamin B12 (µg)	6.85 (3.31)	4.63 (2.16)	<.001	<.001
Vitamin C (mg)	148.2 (77.0)	106.6 (73.1)	.02	.05
Retinol (µg)	426.1 (330.3)	236.2 (137.7)	.001	.001
Vitamin D (µg)	3.47 (2.15)	2.55 (1.61)	.049	.049
Vitamin E (mg)	9.11 (3.36)	7.84 (2.77)	.13	.13
Salt (g)	5.91 (2.7)	6.48 (2.1)	<.001	<.001

[a] MUFA: monounsaturated fatty acids; PUFA: polyunsaturated fatty acids; RE: retinol equivalents; SFA: saturated fatty acids; TE: total energy.

[b] Controlled for gender where appropriate. No significant interactions were observed between method and gender.

[c] *P* value derived using 2-samples paired *t* test.

After controlling for energy, estimated intakes of macronutrients were similar for both the WFR and the Food4Me FFQ with no significant differences between intakes of total fat (g, TE), polyunsaturated fatty acids (PUFA g, %TE), protein (g, %TE), carbohydrate (g), and total sugars (g, %TE) (Table 5). However, estimated intakes of saturated fatty acids (SFA) (g, %TE) and monounsaturated fatty acids (MUFA) (g, %TE) were significantly higher (*P*<.001), and estimated intake of carbohydrate (%TE) was significantly lower (*P*=.01), for the FFQ than for the WFR. For micronutrients, no significant differences were observed between energy-controlled estimates of folate, iron, thiamin, vitamin B6, and vitamin E. Estimated intakes of calcium, total carotene, riboflavin, vitamin B12, vitamin C, retinol, and vitamin D intakes were significantly different between 4-day WFR and FFQ1 (all were higher for the FFQ). After controlling for energy and, where appropriate, gender, vitamin C intakes were no longer significantly different. Removing underreporters from the dataset reduced the agreement between the 2 methods for folate, vitamin B6, and vitamin E, but improved agreement for energy-controlled carbohydrate and vitamin D, with no significant differences observed between estimates of these nutrients (data not shown).

Bland-Altman plots for mean energy (kcal), total fat (%TE), protein (%TE) and carbohydrate (%TE) for the 4-day WFR and FFQ1 are shown in Figure 4. Overall, less than 5% of cases fell outside of the limits of agreement for all plots indicating good agreement between the methods. The mean difference (bias) between energy intakes was relatively small (178 kcal/day) with greater values being estimated in the Food4Me FFQ, as was the case for energy derived from total fat and protein.

Correlation coefficients for estimates of energy and nutrient intakes and cross-classification of quartiles of mean daily intakes between 4-day WFR and FFQ1 are presented in Table 6. Correlation coefficients ranged from .23 (vitamin D) to .65 (protein, %TE) with a mean value of .47. Correlation was significant for the majority of nutrients at the *P*<.01 level, with the exception of total fat (%TE), PUFA (%TE), and vitamin D. Retinol and vitamin E showed significant correlation at the *P*<.05 level. The percentage of participants classified into quartiles of exact agreement ranged from 22% (total fat, %TE) to 53% (MUFA, g). For classifications of exact agreement plus adjacent, values were consistently high, ranging from 65% (sodium) to 88% (total fat, g, and total sugars, g, %TE). The mean percentage of participants classified into quartiles of disagreement was 16% with less than 4% of participants classified into extreme disagreement.

Table 6. Unadjusted correlation coefficients and cross-classification of quartiles of mean energy and nutrient intakes derived from the online Food4Me FFQ and 4-day WFR (n=49).

Nutrient[a]	Correlation	Quartiles, %			
		Exact agreement[b]	Exact agreement plus adjacent[c]	Disagreement[d]	Extreme disagreement[e]
Energy (kcal)	.53[f,h]	41	84	12	2
Total fat (g)	.56[h]	37	88	10	2
Total fat (%TE)	.27	22	76	16	8
SFA (g)	.48[h]	37	82	14	4
SFA (%TE)	.38[f,h]	24	78	14	8
MUFA (g)	.56[h]	53	86	6	8
MUFA (%TE)	.45[h]	47	86	8	6
PUFA (g)	.45[h]	49	76	22	2
PUFA (%TE)	.24	27	71	24	4
Protein (g)	.59[h]	45	84	14	4
Protein (%TE)	.65[f,h]	45	86	14	0
Carbohydrate (g)	.43[f,h]	37	82	10	8
Carbohydrate (%TE)	.59[f,h]	49	82	18	0
Total sugars (g)	.60[f,h]	41	88	10	2
Total sugars (%TE)	.61[f,h]	45	88	10	2
Alcohol (g)	.61[h]	45	80	16	4
Calcium (mg)	.47[f,h]	41	73	17	0
Total folate (µg)	.58[h]	45	86	10	4
Iron (mg)	.50[h]	41	82	14	4
Total carotene (µg)	.42[h]	33	78	18	4
Riboflavin (mg)	.50[h]	45	84	14	2
Thiamin (mg)	.60[h]	43	82	16	2
Vitamin B6 (mg)	.44[h]	37	78	20	2
Vitamin B12 (µg)	.46[h]	39	78	20	2
Vitamin C (mg)	.54[h]	37	84	14	2
Retinol (µg)	.31[g]	37	76	18	6
Vitamin D (µg)	.23	27	67	27	6
Vitamin E (mg)	.30[g]	33	78	14	8
Sodium (mg)	.37[h]	49	65	31	4
Salt (g)	.37[h]	49	65	31	4

[a] MUFA: monounsaturated fatty acids; PUFA: polyunsaturated fatty acids; RE: retinol equivalents; SFA: saturated fatty acids; TE: total energy.

[b] Exact agreement, % of cases cross-classified into the same quartile.

[c] Exact agreement plus adjacent, % of cases cross-classified into the same or adjacent quartile.

[d] Disagreement, % of cases cross-classified 2 quartiles apart.

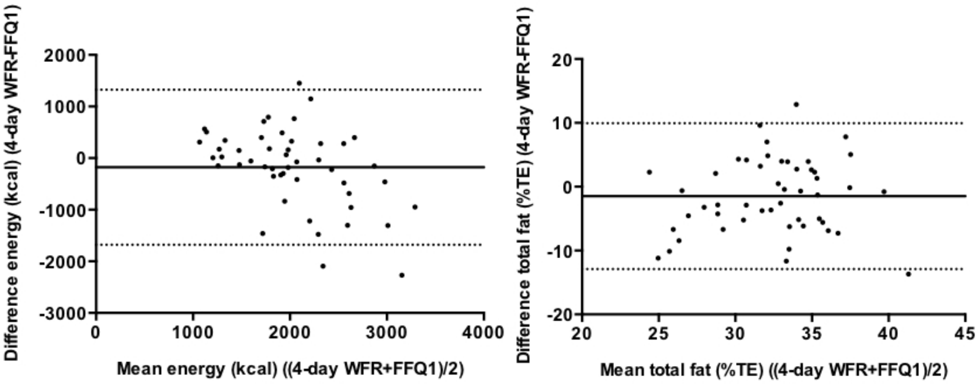
[e] Extreme disagreement, % of cases cross-classified into extreme quartiles.

[f] Pearson correlation.

[g] $P<.05$.

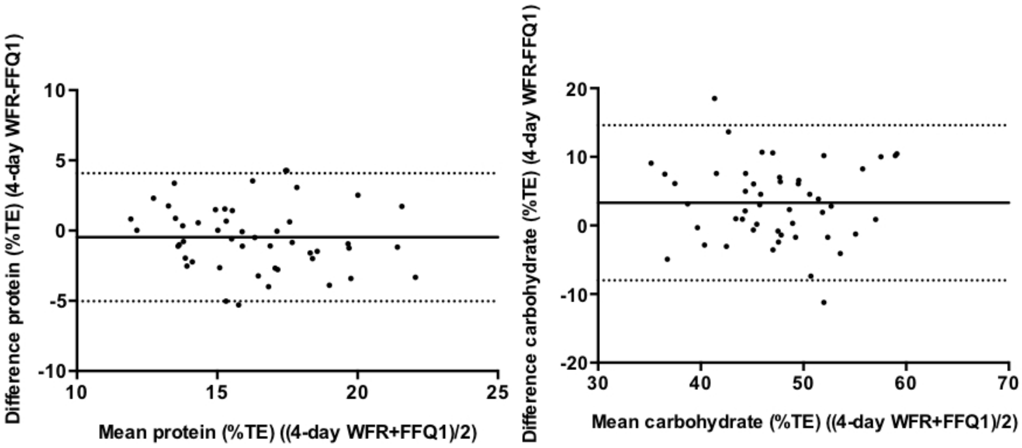
[h] $P<.01$.

Figure 4. Validation study Bland-Altman plots for (a) energy, (b) total fat, (c) protein, and (d) carbohydrate with the mean difference and limits of agreement. The solid line represents the mean difference and the dotted lines represent the limits of agreement.



a. Energy (kcal)

b. Total Fat (%TE)



c. Protein (%TE)

d. Total Carbohydrate (%TE)

Comparison of Food Group Intakes Between the Food4Me FFQ and 4-Day WFR

To assess differences in food group intakes between 4-day WFR and FFQ1, food items were categorized into 35 food groups. Correlation coefficients and cross-classification of mean food group intakes are presented in Table 7. SCC ranged widely from .11 (soups, sauces and miscellaneous foods) to .73 (yogurts) with a mean value of .2. Correlations were significant for the 63% of food groups (22 of 35).

The percentage of participants classified into quartiles of exact agreement ranged from 18% (nuts and seeds, herbs and spices) to 55% (teas and coffees). For classifications of exact agreement plus adjacent, values were high ranging from 55% (soups, sauces and miscellaneous foods) to 90% (milk, chipped, fried and roast potatoes, teas and coffees, and other beverages) with a mean of 78%. The mean percentage of participants classified into quartiles of disagreement was 17% and for extreme disagreement was 5%.

Table 7. Spearman correlation coefficients (SCC) and cross-classification of quartiles of food group intake derived from the online Food4Me FFQ and 4-day WFR (n=49).

Food group	SCC	Quartiles, %			
		Exact agreement[a]	Exact agreement plus adjacent[b]	Disagreement[c]	Extreme disagreement[d]
Rice, pasta, grains and starches	.34[e]	45	73	22	4
Savories (lasagne, pizza)	.16	22	76	22	2
White bread (rolls, tortillas, crackers)	.46[f]	42	73	22	4
Wholemeal and brown breads and rolls	.33[e]	31	65	29	6
Breakfast cereals and porridge	.27	31	78	14	8
Biscuits	.47[f]	41	78	16	6
Cakes, pastries and buns	.31[e]	31	73	18	8
Milk	.68[f]	51	90	8	2
Cheeses	.46[f]	41	82	14	4
Yogurts	.73[f]	43	96	4	0
Ice cream, creams and desserts	.21	35	78	18	4
Eggs and egg dishes	.55[f]	43	76	20	4
Fats and oils (eg, butter, low-fat spreads, hard cooking fats)	.35[e]	35	80	12	8
Potatoes and potato dishes	.38[f]	31	71	20	8
Chipped, fried and roasted potatoes	.52[f]	39	90	8	2
Peas, beans and lentils and vegetable and pulse dishes	.23	24	76	16	8
Green vegetables	.44[f]	37	86	10	4
Carrots	.14	20	82	16	2
Salad vegetables (eg, lettuce)	.23	39	69	18	12
Other vegetables (eg, onions)	.15	41	73	10	16
Tinned fruit or vegetables	.16	29	71	22	6
Bananas	.45[f]	29	80	20	0
Other fruits (eg, apples, pears, oranges)	.47[f]	49	76	22	2
Nuts and seeds, herbs and spices	.23	18	78	14	8
Fish and fish products/dishes	.60[f]	49	82	16	2
Bacon and ham	.53[f]	39	84	16	0
Red meat (eg, beef, veal, lamb, pork)	.26	43	80	16	4
Poultry (chicken and turkey)	.58[f]	45	86	10	4
Meat products (eg, burgers, sausages, pies, processed meats)	.20	27	78	22	0
Alcoholic beverages	.59[f]	43	73	22	4
Sugars, syrups, preserves and sweeteners	.36[f]	37	76	16	8
Confectionary and savory snacks	.25	29	71	22	6
Soups, sauces and miscellaneous foods	.11	24	55	31	14
Teas and coffees	.62[f]	55	90	6	4

Food group	SCC	Quartiles, %			
		Exact agreement[a]	Exact agreement plus adjacent[b]	Disagreement[c]	Extreme disagreement[d]
Other beverages (eg, fruit juices, carbonated beverages, squash)	.66[f]	43	90	8	2

[a] Exact agreement, % of case cross-classified into the same quartile.

[b] Exact agreement plus adjacent, % of cases cross-classified into the same or adjacent quartile.

[c] Disagreement, % of cases cross-classified 2 quartiles apart.

[d] Extreme disagreement, % of cases cross-classified into extreme quartiles.

[e] *P*<.05.

[f] *P*<.01.

Usability Rating

Mean values and standard deviations for responses to the dietary record usability-rating questionnaire are shown in Table 8.

The Food4Me FFQ was considered significantly easier and less time consuming to complete than the 4-day WFR. However, the 4-day WFR was rated as significantly more interesting than the Food4Me FFQ, making participants reflect more on their food intake. Participants were more willing to complete further Food4Me FFQ than 4-day WFR.

Table 8. Responses to Dietary Record Usability-Rating Questionnaire (n=48).

Question	Questionnaire[a], mean (SD)		*P* value[b]
	Food4Me FFQ	4-day WFR	
1. Easy to complete	1.89 (0.71)	2.13 (0.88)	<.001
2. Too time consuming	3.43 (1.09)	3.00 (0.94)	<.001
3. Interesting to complete	2.20 (0.69)	2.07 (0.72)	.006
4. Made me reflect on my intake	2.13 (0.58)	1.89 (0.77)	.002
5. I would be willing to complete more	1.78 (0.70)	2.07 (0.77)	<.001

[a] 1=strongly agree, 2=agree, 3=neither agree nor disagree, 4=disagree, 5=strongly disagree.

[b] *P* value derived using 2-samples paired *t* test.

Discussion

Main Findings and Comparisons With Other Work

Previous validation of the Food4Me FFQ has demonstrated good agreement with the printed EPIC-Norfolk FFQ for the estimation of food and nutrient intake. In this study, participants were asked to complete the Food4Me and EPIC-Norfolk FFQ in a random order, 4 weeks apart. Good agreement between cross-classifications of daily energy and nutrient intakes, estimated using the 2 FFQ, demonstrated the utility of the Food4Me FFQ for ranking individuals based on their nutrient intake. However, it was noted that further testing of the Food4Me FFQ was required to establish its wider utility [18]. The present study thus aimed to demonstrate the reproducibility of the Food4Me FFQ and its validity against a 4-day WFR.

Overall, the Food4Me FFQ demonstrated good reproducibility for the estimation of intakes of nutrients and food groups. Reported energy intakes were significantly lower with the second administration of the Food4Me FFQ, but correlations between energy intakes were high (*r*=.77). Correlation coefficients for nutrient intakes ranged from .65-.90, showing above-average performance compared with the range of .50-.80 proposed by Willet [15]. The mean unadjusted correlation coefficient (*r*=.75) for energy and nutrient intake compared well with previous studies on both computerized [21,30-31] and non-Web-based FFQ [23,32-35]. Associations between food group intakes were similarly strong with an average unadjusted SCC of .75; previous studies have reported correlations of .66 and .72 [36-37]. Cross-classification analysis of repeated measures of intakes of energy, nutrients, and food groups indicated a high level of reproducibility with classification into quartiles of exact agreement plus adjacent averaging 92% for energy and nutrient intake and for food group intakes. Cross-classifications were within the range reported by previous studies [21,32,38]. Bland-Altman plots demonstrated a good level of reproducibility for energy-controlled total fat, protein, and carbohydrate intake, which reinforces evidence for the reliability of the Food4Me FFQ.

Estimated energy and nutrient intakes were higher on the first administration of the Food4Me FFQ than on the second administration. This pattern has been observed in numerous other reproducibility studies [21,31-32], and is proposed to result from learning effects and questionnaire fatigue [39]. The above-average reproducibility of the Food4Me FFQ could be attributed to the addition of food photographs to the FFQ for the estimation of portion size intake. Use of tools that allow participants to report their own portion sizes tend to report higher correlation coefficients between repeat administrations

[19]. The relatively short interval between repeat administrations of the FFQ is another factor likely to have contributed to the questionnaire's good performance. Tsubono et al [40] found that correlation coefficients tended to be lower when FFQ were repeated after a long time interval (6 months to 1 year) compared with a shorter time interval (1 to 6 months), and proposed that the temporal difference may be due to changes in dietary habits which are more likely to occur with longer time intervals. In addition, it has also been suggested that for very short time intervals between administrations, respondents may remember and replicate their entries rather than reporting their diet intake accurately [23]. However, with a large FFQ containing 157 food items, as used in the present study, it is unlikely that many participants would be able to remember their earlier responses.

The degree of underreporting between the Food4Me FFQ and 4-day WFR varied, with 12 (24%) and 15 (15%) participants deemed to be underreporting in the 4-day WFR and Food4Me FFQ, respectively. Given that the WFR is described as the gold standard for assessing intake, our observation that estimates of energy intake were similar between the Food4Me FFQ and the 4-day WFR and that a smaller proportion of participants appeared to underreport with the Food4Me FFQ suggests that the Food4Me FFQ is a promising tool for estimating habitual food intake.

Overall, the results of the validation study showed moderate agreement between the Food4Me FFQ and 4-day WFR for the estimation of energy and nutrient intake. Ranks of energy and nutrient intake estimated using the Food4Me FFQ were highly comparable to the 4-day WFR with the percentage of individuals classified into quartiles of exact agreement and exact agreement plus adjacent averaging 40% and 80%, respectively. Previous studies comparing FFQ with food records have reported average exact agreement classifications between 34% and 49% [32,38,41] and exact agreement plus adjacent quartile classifications of 77% [21]. Estimates of intake showing disagreement between measurement tools in the present study were small and were comparable with the aforementioned studies. Cross-classifications of estimates of food group intake were similar to that of the nutrients, with classification into quartiles of exact agreement plus adjacent averaging 78% and Bland-Altman plots demonstrated good agreement between the 2 methods for estimates of energy and energy-adjusted macronutrient intakes.

In the present study, 22 of 30 nutrients assessed had a correlation coefficient greater than the .4 threshold that was proposed by Cade at al. [23], and 13 of 30 achieved a correlation greater than or equal to the "desirable" .5 proposed by Masson et al [42]. The average unadjusted correlation coefficient of .47 compared favorably with the range reported by similar validation studies comparing FFQ with food records: .34-.46 [31-32,34,41]. SCC for food group intakes were highly variable, ranging from .11 (soups, sauces and miscellaneous foods) to .73 (yogurts), with a mean value of .2. Similar studies have reported correlations ranging from .09 to .83 [37], .17 to .95 [36], and .09 to .58 [41] with mean values of .38, .63, and .58, respectively. However, it is difficult to compare results from these studies because the type of food records and time intervals between dietary assessments differed substantially and there may be substantial differences in the food items included in particular food groups in each of the studies. Variation between Food4Me FFQ and 4-day WFR estimates were greatest for soups, sauces and spreads, carrots, other fruit and vegetables, and tinned fruit and vegetables. It is possible that intakes of these foods might have been overestimated in the Food4Me FFQ, as has been observed previously when several food items within a food group are listed separately in a questionnaire (eg, carrots could be counted under both fresh and tinned carrots and under tinned vegetables) [43]. In addition, foods perceived as healthy, such as fruit and vegetables, are prone to overestimation in FFQ. Furthermore, because they refer to just 4 days' intake, WFR provide a limited snapshot of an individual's diet only and are less able to assess patterns of dietary intake than the Food4Me FFQ, which attempts to capture intakes over the previous month. It is thought that individuals may be able to more accurately estimate the consumption of some foods (eg, alcoholic beverages) than others, as was the case in the present study [21,44]. Alcohol is often considered a confounder in nutrition research given that it constitutes the difference between food and total dietary energy intake; therefore, it is important that it is estimated reliably using the Food4Me FFQ. It is also encouraging that estimates of fish products were well correlated, as these foods are eaten less frequently and may be prone to underrepresentation in 4-day WFR. However, it is surprising that some more commonly consumed foods such as breakfast cereals and porridge show much weaker correlation (*r*=.27).

Our observation that participants in the present study reported that the Food4Me FFQ was easier to use and less time consuming compared with the 4-day WFR, is promising given the movement of health service delivery toward Web-based interventions. Moreover, completion of the Food4Me FFQ was associated with less reflection by participants on their dietary intake, which is known to influence eating behavior. Minimizing the impact of a questionnaire on dietary behavior is beneficial in nutrition intervention studies to ensure that study outcomes are not biased by the methods used for dietary assessment.

Strengths and Limitations

Strengths of the present study include the comparison of the Food4Me FFQ with the gold standard, a WFR, and the use of multiple methods to assess the validation and reproducibility of the Food4Me FFQ. In addition, this validation study had an adequate sample size [15] similar to those used in previous studies [30,45-46]. It should be noted that the validation of the Food4Me FFQ was assessed in a convenient rather than a nationally representative sample of the population, although the inclusion/exclusion criteria used were the same as those intended for the Food4Me study [16]. The use of a convenient university population, with a potentially higher education level, may have implications on the ability of the wider population to complete the online Food4Me FFQ.

Limitations of the study include the use of those recruited first to complete the 4-day WFR because these individuals may have been more motivated to comply with the guidelines. A further limitation is the use of nonconsecutive days in the 4-day WFR, which may have resulted in participants making up food intake on the days they do not record (eg, eating healthy for the record days and overconsuming between record days). However, nonconsecutive recording does have the advantage of capturing a greater diversity of food intake over a week's period.

A potential criticism of the assessment of reproducibility is the short duration between repeated measures. It has been suggested that for very short time intervals, respondents may remember and replicate their entries rather than accurately reporting their dietary intake [23], but this is unlikely to be a significant problem in the present study where the Food4Me FFQ contained 157 food items. Remembering their responses to such a long list of questions after a period of 4 weeks is a memory challenge beyond most people's abilities. The average reported correlation coefficient for crude total fat intake using FFQs repeated after 1 month or less was .68 [19], which compares very favorably with the correlation coefficient of .81 in the present study, showing above-average performance for the Food4Me FFQ. Cade et al [23] suggested that the time interval between repeated measures using a dietary instrument should be chosen to minimize changes in dietary intake and our use of 4 weeks fits that criterion.

Conclusions

In conclusion, the self-administered online Food4Me FFQ demonstrates good reproducibility for the estimation of energy, nutrient, and food group intakes and moderate agreement for the assessment of energy and nutrient intakes when compared with a 4-day WFR in an adult population. Consequently, the online Food4Me FFQ was considered suitable for the assessment of dietary intake in healthy adults.

Acknowledgments

This work was funded by Food4Me (KBBE.2010.2.3-02, Project no 265494). The authors would like to thank Kate Currie and Marzena Kucharska-Kelly for their assistance with the running of the study, Crème Global (Dublin, Ireland) for their support with the online Food4Me FFQ, and Breige McNulty for her assistance with analysis of the WFR.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Food4Me Food Frequency Questionnaire screenshots.

[PDF File (Adobe PDF File), 242KB - jmir_v16i8e190_app1.pdf]

References

1. Rayner M, Scarborough P. The burden of food related ill health in the UK. J Epidemiol Community Health 2005 Dec;59(12):1054-1057 [FREE Full text] [doi: 10.1136/jech.2005.036491] [Medline: 16286493]
2. Scarborough P, Bhatnagar P, Wickramasinghe KK, Allender S, Foster C, Rayner M. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs. J Public Health (Oxf) 2011 Dec;33(4):527-535 [FREE Full text] [doi: 10.1093/pubmed/fdr033] [Medline: 21562029]
3. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, et al. Program on the Global Demography of Aging. Geneva: World Economic Forum; 2011. The global economic burden of noncommunicable diseases URL: http://www.hsph.harvard.edu/program-on-the-global-demography-of-aging/WorkingPapers/2012/PGDA_WP_87.pdf [accessed 2014-02-25] [WebCite Cache ID 6RfkcHGd4]
4. Miniwatts Marketing Group. World Internet Usage and Population Statistics. Colombia: Miniwatts Marketing Group 2012 URL: http://www.internetworldstats.com/stats.htm [accessed 2014-08-06] [WebCite Cache ID 6RctWnShh]
5. Kolasa KM, Guy Miller M. New developments in nutrition education using computer technology. Journal of Nutrition Education 1996 Jan;28(1):7-14. [doi: 10.1016/S0022-3182(96)70009-2]
6. Sacks N, Cabral H, Kazis LE, Jarrett KM, Vetter D, Richmond R, et al. A web-based nutrition program reduces health care costs in employees with cardiac risk factors: before and after cost analysis. J Med Internet Res 2009;11(4):e43 [FREE Full text] [doi: 10.2196/jmir.1263] [Medline: 19861297]
7. Solomon M, Wagner SL, Goes J. Effects of a Web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012;14(1):e32 [FREE Full text] [doi: 10.2196/jmir.1924] [Medline: 22353433]
8. Kodama S, Saito K, Tanaka S, Horikawa C, Fujiwara K, Hirasawa R, et al. Effect of Web-based lifestyle modification on weight control: a meta-analysis. Int J Obes (Lond) 2012 May;36(5):675-685. [doi: 10.1038/ijo.2011.121] [Medline: 21694698]
9. Adamson AJ, Mathers JC. Effecting dietary change. Proc Nutr Soc 2004 Nov;63(4):537-547. [Medline: 15831125]
10. Coulston AM, Boushey CJ. Nutrition in the Prevention and Treatment of Disease. Amsterdam: Academic Press; 2008.
11. Rebro SM, Patterson RE, Kristal AR, Cheney CL. The effect of keeping food records on eating patterns. J Am Diet Assoc 1998 Oct;98(10):1163-1165. [doi: 10.1016/S0002-8223(98)00269-7] [Medline: 9787724]

12. Block G. A review of validations of dietary assessment methods. Am J Epidemiol 1982 Apr;115(4):492-505. [Medline: 7041631]
13. Livingstone MB, Prentice AM, Strain JJ, Coward WA, Black AE, Barker ME, et al. Accuracy of weighed dietary records in studies of diet and health. BMJ 1990 Mar 17;300(6726):708-712 [FREE Full text] [Medline: 2386561]
14. Matthys C, Pynaert I, De Keyzer W, De Henauw S. Validity and reproducibility of an adolescent web-based food frequency questionnaire. J Am Diet Assoc 2007 Apr;107(4):605-610. [doi: 10.1016/j.jada.2007.01.005] [Medline: 17383266]
15. Willett W. Nutritional Epidemiology. New York: Oxford University Press; 1998.
16. Food4Me. 2011. URL: http://www.food4me.org [WebCite Cache ID 6J5McFpZv]
17. Walsh M, Kuhn S. Developments in personalised nutrition. Nutrition Bulletin 2012 Nov 08;37(4):380-383 [FREE Full text] [doi: 10.1111/j.1467-3010.2012.01998.x]
18. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: the Food4Me food frequency questionnaire. J Med Internet Res 2014;16(6):e150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]
19. Cade JE, Burley VJ, Warm DL, Thompson RL, Margetts BM. Food-frequency questionnaires: a review of their design, validation and utilisation. Nutr Res Rev 2004 Jun;17(1):5-22. [doi: 10.1079/NRR200370] [Medline: 19079912]
20. Swierk M, Williams PG, Wilcox J, Russell KG, Meyer BJ. Validation of an Australian electronic food frequency questionnaire to measure polyunsaturated fatty acid intake. Nutrition 2011 Jun;27(6):641-646. [doi: 10.1016/j.nut.2010.06.011] [Medline: 20869207]
21. Labonté M, Cyr A, Baril-Gravel L, Royer MM, Lamarche B. Validity and reproducibility of a web-based, self-administered food frequency questionnaire. Eur J Clin Nutr 2012 Feb;66(2):166-173. [doi: 10.1038/ejcn.2011.163] [Medline: 21934698]
22. Christensen SE, Möller E, Bonn SE, Ploner A, Wright A, Sjölander A, et al. Two new meal- and web-based interactive food frequency questionnaires: validation of energy and macronutrient intake. J Med Internet Res 2013;15(6):e109 [FREE Full text] [doi: 10.2196/jmir.2458] [Medline: 23739995]
23. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr 2002 Aug;5(4):567-587. [doi: 10.1079/PHN2001318] [Medline: 12186666]
24. Nelson M, Atkinson M, Meyer J. Food Portion Sizes: A Photographic Atlas. London: MAFF Publications; 1997.
25. 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]
26. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr 1991 Dec;45(12):569-581. [Medline: 1810719]
27. Hébert JR, Peterson KE, Hurley TG, Stoddard AM, Cohen N, Field AE, et al. The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees. Ann Epidemiol 2001 Aug;11(6):417-427. [Medline: 11454501]
28. Tinuviel Software. 2003. Food composition databank URL: http://www.tinuvielsoftware.com/wisp41.htm [accessed 2013-08-23] [WebCite Cache ID 6J5NOm32z]
29. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
30. Engle A, Lynn LL, Koury K, Boyar AP. Reproducibility and comparability of a computerized, self-administered food frequency questionnaire. Nutr Cancer 1990;13(4):281-292. [doi: 10.1080/01635589009514070] [Medline: 2345706]
31. Beasley JM, Davis A, Riley WT. Evaluation of a web-based, pictorial diet history questionnaire. Public Health Nutr 2009 May;12(5):651-659 [FREE Full text] [doi: 10.1017/S1368980008002668] [Medline: 18547450]
32. Goulet J, Nadeau G, Lapointe A, Lamarche B, Lemieux S. Validity and reproducibility of an interviewer-administered food frequency questionnaire for healthy French-Canadian men and women. Nutr J 2004 Sep 13;3:13 [FREE Full text] [doi: 10.1186/1475-2891-3-13] [Medline: 15363100]
33. Boucher B, Cotterchio M, Kreiger N, Nadalin V, Block T, Block G. Validity and reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. Public Health Nutr 2006 Feb;9(1):84-93. [Medline: 16480538]
34. Collins CE, Boggess MM, Watson JF, Guest M, Duncanson K, Pezdirc K, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian adults. Clin Nutr 2013 Oct 9. [doi: 10.1016/j.clnu.2013.09.015] [Medline: 24144913]
35. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992 May 15;135(10):1114-1126. [Medline: 1632423]
36. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. J Am Diet Assoc 1993 Jul;93(7):790-796. [Medline: 8320406]
37. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, et al. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am J Clin Nutr 1999 Feb;69(2):243-249 [FREE Full text] [Medline: 9989687]

38. Chiu TH, Huang HY, Chen KJ, Wu YR, Chiu JP, Li YH, et al. Relative validity and reproducibility of a quantitative FFQ for assessing nutrient intakes of vegetarians in Taiwan. Public Health Nutr 2014 Jul;17(7):1459-1466. [doi: 10.1017/S1368980013001560] [Medline: 23782894]
39. Wong SS, Boushey CJ, Novotny R, Gustafson DR. Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and non-Hispanic white youth. J Am Diet Assoc 2008 Mar;108(3):539-543. [doi: 10.1016/j.jada.2007.12.006] [Medline: 18313437]
40. Tsubono Y, Nishino Y, Fukao A, Hisamichi S, Tsugane S. Temporal change in the reproducibility of a self-administered food frequency questionnaire. Am J Epidemiol 1995 Dec 1;142(11):1231-1235. [Medline: 7485070]
41. Torheim LE, Ouattara F, Diarra MM, Thiam FD, Barikmo I, Hatløy A, et al. Nutrient adequacy and dietary diversity in rural Mali: association and determinants. Eur J Clin Nutr 2004 Apr;58(4):594-604. [doi: 10.1038/sj.ejcn.1601853] [Medline: 15042127]
42. Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. Public Health Nutr 2003 May;6(3):313-321. [doi: 10.1079/PHN2002429] [Medline: 12740081]
43. Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J. Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol 1997;26 Suppl 1:S59-S70 [FREE Full text] [Medline: 9126534]
44. Fidanza F, Gentile MG, Porrini M. A self-administered semiquantitative food-frequency questionnaire with optical reading and its concurrent validation. Eur J Epidemiol 1995 Apr;11(2):163-170. [Medline: 7672070]
45. Munger RG, Folsom AR, Kushi LH, Kaye SA, Sellers TA. Dietary assessment of older Iowa women with a food frequency questionnaire: nutrient intake, reproducibility, and comparison with 24-hour dietary recall interviews. Am J Epidemiol 1992 Jul 15;136(2):192-200. [Medline: 1415141]
46. Riley MD, Blizzard L. Comparative validity of a food frequency questionnaire for adults with IDDM. Diabetes Care 1995 Sep;18(9):1249-1254. [Medline: 8612438]

Abbreviations

BMI: body mass index
EPIC: European Prospective Investigation of Cancer
FFQ: food frequency questionnaire
MUFA: monounsaturated fatty acids
PUFA: polyunsaturated fatty acids
RE: retinol equivalents
SFA: saturated fatty acids
TE: total energy
WFR: weighed food record

Edited by G Eysenbach; submitted 27.02.14; peer-reviewed by A Lennox, S Fulton; comments to author 07.07.14; accepted 22.07.14; published 11.08.14.

Please cite as:
Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, Gibney ER, Gibney MJ, Lovegrove JA
Online Dietary Intake Estimation: Reproducibility and Validity of the Food4Me Food Frequency Questionnaire Against a 4-Day Weighed Food Record
J Med Internet Res 2014;16(8):e190
URL: http://www.jmir.org/2014/8/e190/
doi:10.2196/jmir.3355
PMID:25113936

Review

Massive Open Online Courses on Health and Medicine: Review

Tharindu Rekha Liyanagunawardena[1*], BSc (Hons), MBA, PhD; Shirley Ann Williams[1*], BSc, FBCS, CEng, CITP, CSci, FHEA, PhD

School of Systems Engineering, University of Reading, Reading, United Kingdom

[*] all authors contributed equally

Corresponding Author:
Tharindu Rekha Liyanagunawardena, BSc (Hons), MBA, PhD
School of Systems Engineering
University of Reading
Whiteknights
Reading, RG6 6AY
United Kingdom
Phone: 44 118 378 6423
Fax: 44 118 975 1822
Email: t.r.liyanagunawardena@reading.ac.uk

Related Article:

Letter: http://www.jmir.org/2014/8/e197/

Abstract

Background: Massive open online courses (MOOCs) have become immensely popular in a short span of time. However, there is very little research exploring MOOCs in the discipline of health and medicine.

Objective: We aim to provide a review of MOOCs related to health and medicine offered by various MOOC platforms in 2013, by analyzing and comparing the various offerings, their target audience, typical length of course, and credentials offered. We also discuss opportunities and challenges presented by MOOCs in health and medicine.

Methods: Health and medicine–related MOOCs were gathered using several methods to ensure the richness and completeness of data. Identified MOOC platform websites were used to gather the lists of offerings. In parallel, these MOOC platforms were contacted to access official data on their offerings. Two MOOC aggregator sites (Class Central and MOOC List) were also consulted to gather data on MOOC offerings. Eligibility criteria were defined to concentrate on the courses that were offered in 2013 and primarily on the subject of health and medicine. All language translations in this paper were done using Google Translate.

Results: The search identified 225 courses, of which 98 were eligible for the review. Over half (58%, 57/98) of the MOOCs considered were offered on the Coursera platform, and 94% (92/98) of all the MOOCs were offered in English. Universities offered 90 MOOCs, and the John Hopkins University offered the largest number of MOOCs (12/90). Only three MOOCs were offered by developing countries (China, West Indies, and Saudi Arabia). The duration of MOOCs varied from 3-20 weeks with an average length of 6.7 weeks. On average, MOOCs expected a participant to work on the material for 4.2 hours a week. Verified certificates were offered by 14 MOOCs, while three others offered other professional recognition.

Conclusions: The review presents evidence to suggest that MOOCs can be used as a way to provide continuous medical education. It also shows the potential of MOOCs as a means of increasing health literacy among the public.

(J Med Internet Res 2014;16(8):e191) doi:10.2196/jmir.3439

KEYWORDS

eLearning; education; health education; continuing education; computer-assisted instruction

Introduction

Background

Massive open online courses (MOOCs) are a recent innovative addition to the online learning landscape. They are online courses that are accessible through the Web and open to registration generally without limits on numbers or prerequisites. The course registration and course materials are free of charge, although in some courses one can pay to obtain a certificate of participation or verified certificate (for credit). These courses have start and end dates, but even after the start date, registration is often kept open unlike traditional online courses that close registration at the start of the course. MOOCs carry great potential to reach large numbers of learners from across the world as they can be accessed by anyone anywhere in the world as long as they have Internet access, computer literacy, and language proficiency.

Brief History

"Connectivism and Connective Knowledge" (CCK08), an online course facilitated by George Siemens and Stephen Downes, in 2008, offered through the Learning Technologies Centre and Extended Education at the University of Manitoba [1], is considered the first MOOC [2]. This online course had around 2200 non-credit, non-fee paying students along with 25 paid enrollments (for credit). Unlike traditional online courses that rely primarily on resources posted by the facilitators through a learning management system, this course was conducted according to the principles of connectivism [3], encouraging learning through a network (peer learning) across multiple learning spaces.

Within a short time, MOOCs have attracted wide interest from educators, learners, businesses, media, and the general public. Many prominent universities are now offering their courses as MOOCs. For example, Massachusetts Institute of Technology, Harvard University, Berkeley University of California, and the University of Texas offer MOOCs through the MOOC platform edX. There have also been for-profit ventures such as the Coursera MOOC platform, which partners with over 100 institutions (108 as of March 19, 2014) from around the world.

In some MOOCs, there are hundreds of thousands of enrollments. However, not all students enrolled return for the course and only a small number of them finish all parts of the course [2]. Given the nature of the courses, where participation is voluntary and no financial commitments are made up front, is the number of students who complete the course a concern? If the aim of a MOOC is to provide the opportunity or access to learn from high-quality courses (taught by the experts in the field from world class universities), then the numbers completing the course should not be of prime concern [4]. On the other hand, if the aim were to get everyone registered through to the end, similar to a traditional higher educational institution where a student failing to complete within a given timeframe could adversely affect the university's profile, family, student, and lecturers [5], these completion rates would be a disaster. The problem here could be the use of traditional metrics in this non-traditional or disruptive form of educational provision. However, more evidence-based research may reveal the true nature of MOOCs and possibly better ways of understanding and evaluating them.

Although the MOOC revolution began in North America, it has now spread to universities and institutions in many parts of the developed world. For example, in 2013, the UK MOOC platform, FutureLearn, started offering courses. Initially MOOCs were offered in English, but today there are many MOOCs offered in various languages including Chinese, Arabic, Spanish, and French. For example, in 2012, a Spanish MOOC platform Miriada X was founded, and in 2013, the platform Rawq started offering courses in Arabic. Similarly, XuetangX was created to offer courses in Chinese. However, English remains the dominant language in MOOC provision.

Pedagogy

MOOCs, like other online courses, use a variety of learning materials including videos, documents, and quizzes. At present, MOOCs are mainly classified according to their pedagogical position: connectivist MOOCs (cMOOCs) and "MOOC as eXtension of something else" (xMOOCs) [6]. cMOOCs harness the strength of networks and peer learning generally using multiple learning spaces. Participants in cMOOCs are likely to find a lot of emphasis on participants' stories and learning from them (eg, Rhizomatic Learning: The community is the curriculum on P2PUniversity) than on the learning materials provided by the instructor or course designer. On the other hand, xMOOCs seems to have a more individualist learning approach [7] surrounding the course on a given MOOC platform. In xMOOCs, learning and understanding the content provided in the course is given priority. Original cMOOCs were based on open education practices making their content available using open licensing [8]. However, many xMOOCs offered in platforms such as Coursera use copyrighted materials. However, it is worthwhile noting that there is a continuum of possibilities between these two distinct pedagogical positions.

MOOCs are offered in a wide range of subjects varying from cell biology to astronomy. In this paper, we explore the courses offered by major MOOC platforms on topics related to health and medicine. Several methods were used to collect relevant courses for the review: directly making contact with MOOC platforms to get course data, accessing publicly available information on MOOC platform websites, and using MOOC aggregator sites. Data related to courses offered in 2013 that were collected as earlier offerings lacked relevant details. The paper provides a comprehensive review of MOOCs offered in 2013 in "Health and Medicine" or a related category.

Methods

Data Collection

Overview

In general, researchers use different methods to identify data to be included in a review. For example, to collect papers (data) for a systematic review of literature, researchers would search in databases and/or search engines and chaining from known sources [2]. Similarly, in collecting details of MOOCs offered in topics related to health and medicine for this review, it was important to collect as complete a set of data as possible. A list of MOOCs offered by various providers was not readily available for analysis. Thus, in identifying relevant MOOCs, a range of methods were used to obtain related information that would form a more complete dataset for the analysis.

Platforms

With the growing popularity of MOOCs, there have been various commercial and non-commercial organizations providing platforms where MOOCs can be offered. Identification of such MOOC platforms was carried out using the literature, news items, and Web resources. LISTedTECH (a database of educational companies, educational products, and educational institutions that anyone can edit) lists 19 systems as MOOC platforms as of December 19, 2013 [9]. Using news articles, blogs, and other literature, nine additional MOOC platforms that are in operation were identified. The total of 28 identified MOOC platforms (see Multimedia Appendix 1) and their offerings were considered in this review.

From December 17-21, 2013, each of these MOOC platforms' websites were accessed to find the list of MOOCs offered by each of them on topics relating to health and medicine. In instances where the websites were in languages other than English, Google Translate was used.

Official Records

In parallel, MOOC platform providers were contacted via email to obtain official records when their websites did not have the necessary information. Only five MOOC platform providers (Canvas, iversity, Openlearning, Miriada X, and Crypt4you) responded to this request with information while another MOOC provider (Coursera) responded without the information.

Aggregators

The two MOOC aggregator sites Class Central [10] and MOOC List [11] were also consulted to collect a list of MOOCs.

Eligibility

Free Courses

When platforms provided both paid-for and free courses (such as Udemy), only free courses were considered. Courses offered by University of Miami Global Academy required a US$90 non-refundable one-time registration fee upfront and a tuition fee depending on the number of credits taken. Thus, courses offered by this platform were not included in this review.

Subject

MOOCs listed under "Health and Medicine" or a related category (such as Health Sciences on Miriada X, Health Science on CourseSites, Health and Society on Coursera) were considered. When MOOCs were not categorized (such as OpenupEd and FutureLearn), the course title and where available the course description were used to determine if it was related to health and medicine (eg, "Improving your image: Dental Photography in Practice" on FutureLearn).

MOOCs categorized under Psychology or Biology and Life Sciences (or were predominantly on them) were not considered in this analysis. MOOCs on veterinary sciences but categorized under Health and Medicine (eg, "Canine Theriogenology for Dog Enthusiasts" on Coursera) were also excluded. But when the courses discussed animal health or diseases and their impact/influence on human health, such as "Enfermedades transfronterizas de los animales" (Animal transboundary diseases) on Miriada X, they were included.

Start Date

The time period for the review was defined as January 1 to December 31, 2013 (inclusive). MOOCs having a start date within this period were considered for the review. Self-paced MOOCs (that do not have a specified start date) were omitted. This included 39 courses listed under Health and Medicine in the Veduca platform, 10 courses listed under Health and Fitness in the Udemy platform, and 44 courses listed under Health Literacy on the ALISON platform, and four OpenupEd courses ("Stress post-traumatic disorder: difficulties and debate in making a diagnosis", "Valutazione clinica e strumenti di indagine nell'area traumatica"[Clinical assessment and survey instruments in traumatic area], "Programmi e modelli di intervento nelle situazioni traumatiche" [Programs and intervention models in traumatic situations], and "Anatomo-physiological bases of mental activity"). On the Saylor platform, all courses are self-paced (the titles that seemed relevant were categorized under Biology). The course "La Seguridad del Paciente" (Patient Safety) on Miriada X had to be discounted because the start date for the course could not be established.

Class Central

We found 113 MOOCs related to health and medicine listed in the MOOC aggregator site Class Central [10]. Under the "Finished Courses" section, exactly 100 courses were listed, while 13 were listed in the "Courses in Progress" section (December 24, 2013). A number of courses were excluded for a variety of reasons.

Five courses were excluded from "Courses in Progress":

- "Exploring anatomy: the human abdomen" offered by the University of Leeds on the FutureLearn platform had an incorrect start date in 2013 instead of the correct start date in 2014
- Three self-paced MOOCs ("The Basics of Exercise Programs for Older Adults" on CourseSites, two Stanford University offerings "Practical tips to improve Asian American participation in cancer clinical trials", and "Antimicrobial Stewardship: Optimization of Antibiotic Practices", each 104 weeks long)
- "DEV: Water, Civilization, and Nature: Addressing 21st Century Water Challenges" on CourseSites, which was a self-paced course lacking relevance

We excluded 19 courses from "Finished Courses":

- Nine courses offered in 2012
- Two courses without start dates ("Cardiac Arrest, Hypothermia, and Resuscitation Science" and "Basic Behavioral Neurology" offered by University of Pennsylvania on Coursera).

- Eight courses lacking relevance: “Marathon Training” and “Safety Function & Action: Strategies for Disaster Responders” on Canvas.net, “Critical Thinking in Global Changes” offered by the University of Edinburgh, “Canine Theriogenology for Dog Enthusiasts” offered by University of Minnesota (2 instances), “Equine Nutrition” offered by University of Edinburgh, “Growing Old Around the Globe” offered by the University of Pennsylvania, and “Disaster Preparedness” offered by the University of Pittsburgh on Coursera

The “Understanding Dementia” MOOC was offered by University of Tasmania on Desire2Learn. Desire2Learn was not listed as a MOOC platform as it offered only proprietary software for institutions. But the MOOC was included in the review. Thus a list of 89 relevant MOOCs (out of 113 identified) was obtained from the Class Central aggregator site.

MOOC List

Another MOOC aggregator site, MOOC List [11], listed details of 54 MOOCs in 2013 under “Health and Society” and 45 under “Medicine and Pharmacology” (January 3, 2014). Due to 19 overlapping courses in the two categories, the distinct course count was 80. Out of these, 53 courses overlapped with the list obtained through Class Central, which left a list of 27 new courses. We further disregarded some courses:

- Four self-paced courses: “Bioethics” and “Make the Strategic Case for Disability in the Workplace” on Canvas; “Clinical Psychology” on Saylor and “Enseñanza en consulta y medio hospitalario” (Education in consultation and hospital environment) on CourseSites
- “Introduction to Pharmaceutical Manufacturing” offered by Dublin Institute of Technology on CourseSites with a November 25 start date could not be validated against the MOOC list available from the official website (there was a MOOC “So you want to work in the pharmaceutical industry?... Next Steps” offered by Dublin Institute of Technology on CourseSites and authors believe this entry was thus erroneous)
- 13 courses lacking relevance

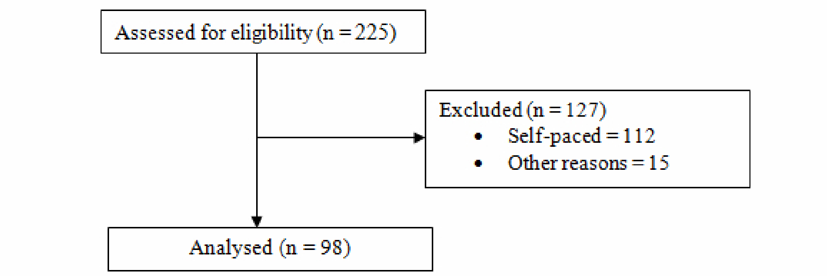
This led to nine entries (seven Coursera courses, a P2P University course, and a course offered by Stanford University VentureLab) from MOOC List being added to the Class Central list (of 89 entries). Therefore, the total number of MOOCs considered for this review is 98 (see Figure 1). The collection of MOOCs included in the review is given in Multimedia Appendix 2. The number of MOOCs from each platform considered in this review is given in Table 1.

Table 1. Number of MOOCs included in the review per platform.

Platform		Total found (N=225)	Self-paced (n=112)	Excluded other reasons (n=15)	Considered for review (n=98)	Not considered for review (n=127)
1	ALISON	44	44		0	44
2	Canvas.net	9	2	2	5	4
3	Coursera	67	2	8	57	10
4	CourseSites	11	3	1	7	4
5	Coursolve	0			0	
6	Crypt4you	0			0	
7	edX	7		2	5[a]	2
8	France Universite Numerique	0			0	
9	FutureLearn	2		1	1	1
10	Galileo Education Systems	0			0	
11	Rwaq	1			1	
12	Iversity	0			0	
13	Miriada X	5		1	4	1
14	NovoEd	1			1	
15	OpenLearning	1	1		0	1
16	Open2Study	14			14	
17	OpenHPI	0			0	
18	OpenupEd	4	4		0	4
19	P2PUniversity	1			1	
20	Saylor	7	7		0	7
21	Skynet	0			0	
22	Udacity	0			0	
23	Udemy	10	10		0	10
24	uneopen.com	0			0	
25	UKeU (not in operation)	0			0	
26	University of Miami Global	0			0	
27	Veduca	39	39		0	39
28	XuetangX	0			0	
29	Stanford University VentureLab	1			1	
30	University of Tasmania on Desire2Learn	1			1	

[a] A course offered by Stanford University in OpenEdX was also considered as edX.

Figure 1. Flow diagram.



Results

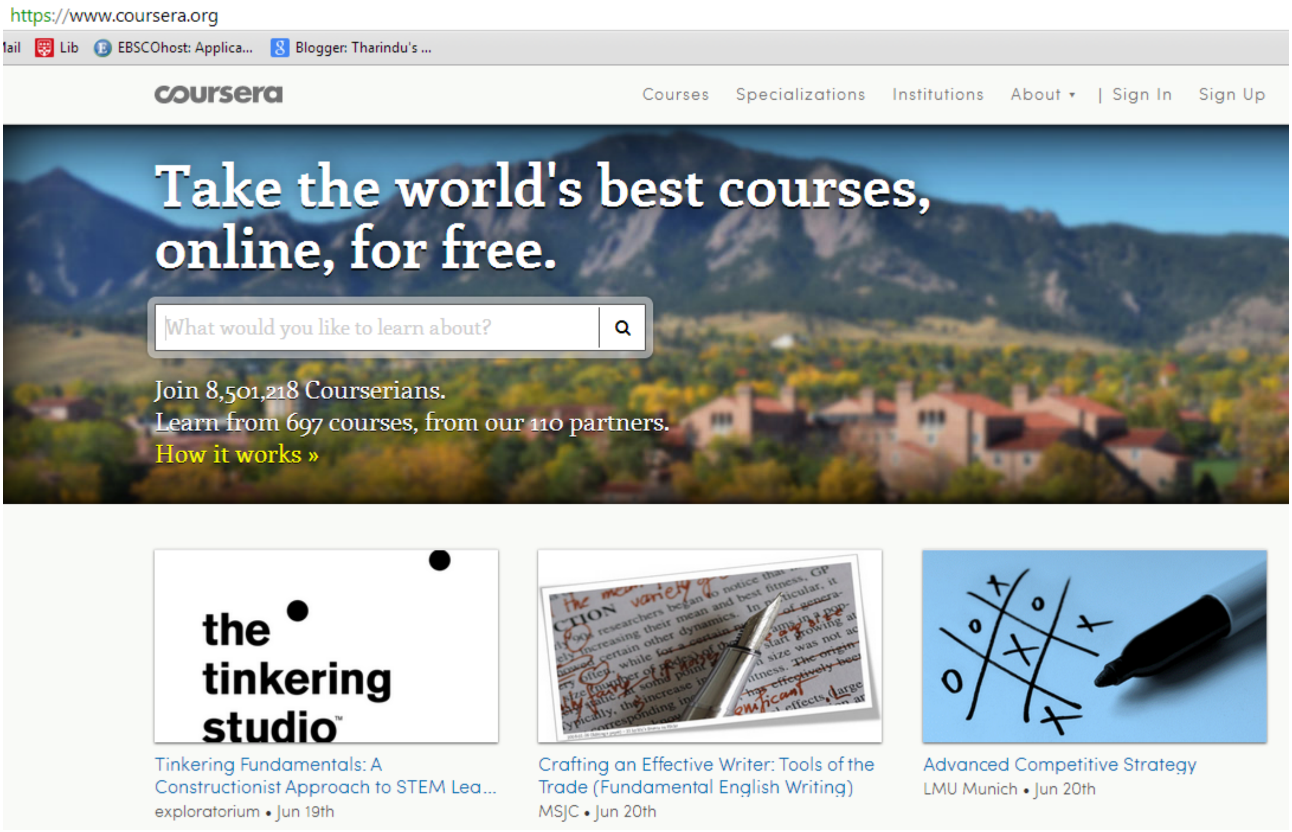
Massive Open Online Course Platforms

Quantitative and qualitative analyses of the data were performed using Microsoft Excel and NVivo software. The majority (58%, 57/98) of MOOCs related to health and medicine was offered by Coursera (Figure 2) followed by Open2Study (Figure 3). Full analysis of course offerings by platform/provider is shown in Table 2.

Table 2. MOOCs by platform/provider (n=98).

Platform/provider	n	%
Coursera	57	58
Open2Study	14	14
CourseSites	7	7
Canvas	5	5
edX	5	5
Miriada X	4	4
FutureLearn	1	1
NovoEd	1	1
P2PUniversity	1	1
Rwaq	1	1
University of Tasmania	1	1
VentureLab	1	1

Figure 2. Coursera platform.



Take the world's best courses, online, for free.

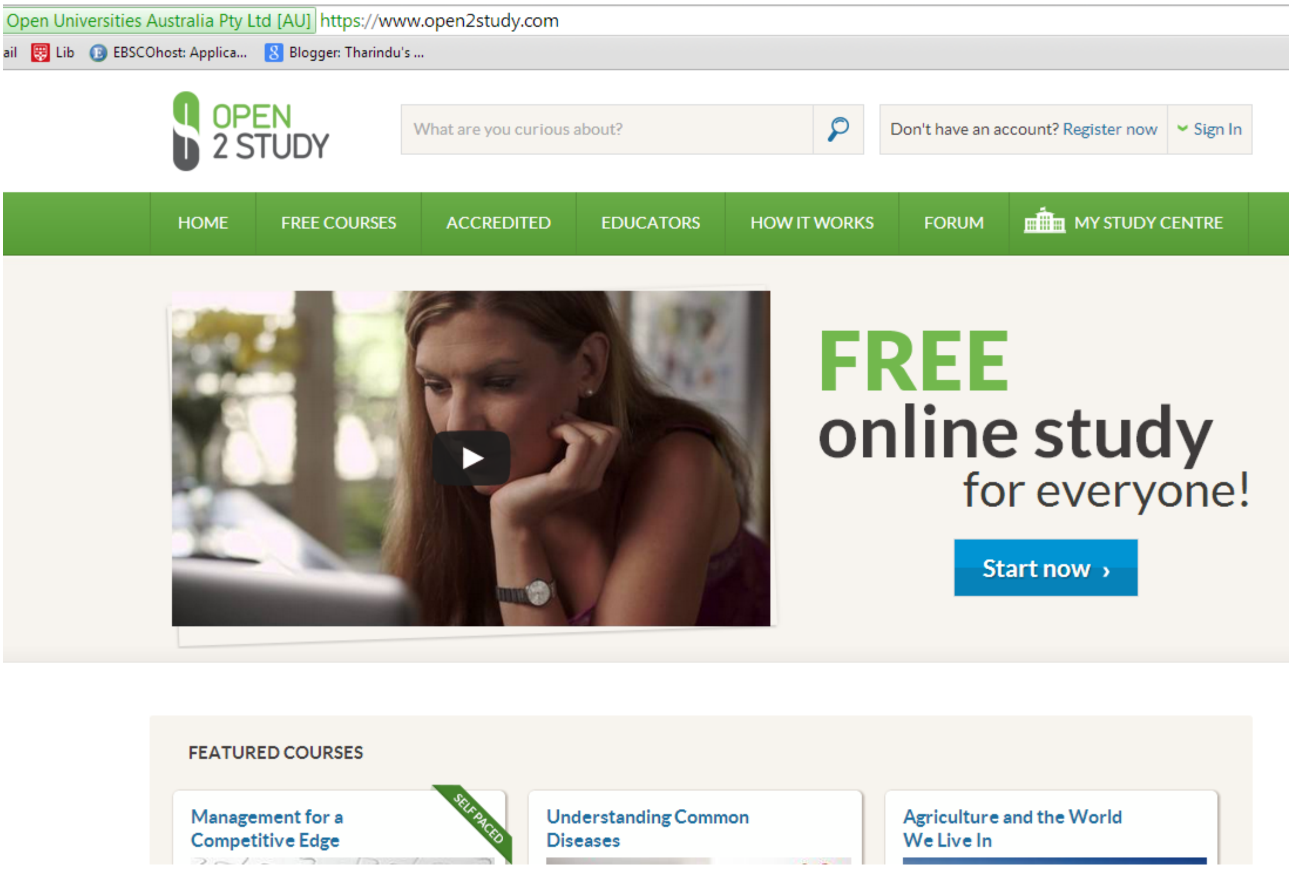
Join 8,501,218 Courserians.
Learn from 697 courses, from our 110 partners.

Tinkering Fundamentals: A Constructionist Approach to STEM Lea...
exploratorium • Jun 19th

Crafting an Effective Writer: Tools of the Trade (Fundamental English Writing)
MSJC • Jun 20th

Advanced Competitive Strategy
LMU Munich • Jun 20th

Figure 3. Open2Study platform.



Language

The language breakdown of MOOCs related to health and medicine shows that the vast majority of MOOCs, 94% (92/98), were offered in English (Table 3). Four MOOCs were offered in Spanish (Castilian) on Miriada X, while one MOOC each was offered in Arabic on Rwaq and in Chinese on Coursera.

Table 3. MOOCs by language (n=98).

Language	n	%
English	92	94
Spanish	4	4
Arabic	1	1
Chinese	1	1

Offering Institution

The highest number of MOOCs in this review were offered by John Hopkins University (12) followed by University of California (nine), University of Pennsylvania (seven), and Open Universities Australia (six). Harvard University and the University of Sheffield offered three MOOCs each.

Most of the MOOCs (90/98) in the review were offered by universities. The large majority of these MOOCs, 70% (63/90) were offered by North American universities. Out of these, only two MOOCs were offered by Canadian universities (University of Toronto). Universities in the other parts of the world offered only a small number of MOOCs in health and medicine (Table 4). These MOOCs were offered by 14 universities (Table 5): five in Australia, four in Spain, two in the United Kingdom, one in each of the Republic of Ireland, Denmark, Switzerland, the West Indies, and China. Also considering the “Introduction to Psychiatry” MOOC offered on Rwaq, very few MOOCs (3/98, 3%) were from developing countries.

Table 4. MOOCs by North American universities versus worldwide (n=90).

	n	%
North America	63	70
Other	27	30

Table 5. MOOCs offered by universities outside North America (n=27).

University	MOOCs, n
Open Universities Australia	6
Dublin Institute of Technology	3
University of Sheffield	3
University of Copenhagen	2
Flinders University	2
The University of Melbourne	1
University of Geneva	1
Shanghai Jiao Tong University	1
University of Birmingham	1
Universidad De Murcia	1
Universidad De Cantabria	1
Universitat Plitecnica De Valencia	1
CEU Universidad San Pablo	1
University of Wollongong	1
University of Tasmania	1
St. George’s University, Grenada	1

Number of Instances

Many MOOCs have run only one instance within the considered period. However, the MOOCs “Food, Nutrition and Your Health” and “Introduction to Nursing in Healthcare” both offered by Open2Study have both run six times. The MOOCs that were offered more than once are listed in Table 6.

Table 6. MOOCs offered multiple times.

MOOC	Platform	Instances, n
Food, Nutrition and Your Health	Open2Study	6
Introduction to Nursing in Healthcare	Open2Study	6
Health for All through Primary Care	Coursera	3
Contraception: Choices, Culture and Consequences	Coursera	2
The Social Context of Mental Health and Illness	Coursera	2
Genes and the Human Condition (From Behavior to Biotechnology)	Coursera	2
Nutrition for Health Promotion and Disease Prevention	Coursera	2
Health Informatics in the Cloud	Coursera	2
So you want to work in the Pharmaceutical Industry	CourseSites	2

Duration

The length of the MOOCs considered for the review varied from 3 weeks (“Introduction to Pharmaceutical Manufacturing Technologies” and “So you want to work in the Pharmaceutical Industry”, two instances) to 20 weeks (“International Health Systems”) with a mode of 6 weeks (21 MOOCs) and average length of 6.7 weeks. In calculating the duration of MOOCs, only 96 MOOCs were considered as the duration of two MOOCs could not be verified. Many MOOCs (75) were 8 weeks or less in duration.

Time Commitment

Most MOOC descriptions (76) contained information on the average time a participant was expected to work on the materials. On average, the MOOCs expected a participant to work on the material for 4.2 hours a week. The Stanford University offering “HRP258: Statistics in Medicine” expected the highest commitment of 8-12 hours per week. Most courses (mode) expected 2-4 hours per week on the course.

Recognition

Some of the MOOCs considered in the review provided certificates for successful participants. The terminology used in different platforms to refer to certificates varied. For example, in Coursera, a “statement of accomplishment” referred to the free certificate signed by the instructor or educator (professor), while in edX a similar credential was referred to as an “honor code certificate”. On Mirianda X, the free certificate was referred to as “certificados de participación” (certificate of participation) and the paid-for certificate was referred to as “certificado de superación” (certificate of overcoming).

According to course descriptions, the Stanford University course “HRP258: Statistics in Medicine” offered a certificate of participation to students who obtained 60% or higher and a certificate with distinction for participants obtaining 90% or higher. Some MOOC descriptions specifically mentioned that the awarded certificate does not carry credits. These were offered by Emory University, the University of Melbourne, University of Geneva, the College of St. Scholastica, and the University of Pennsylvania.

Verified certificates were offered for some courses while some others were eligible for continuous professional development credit. John Hopkins University offered four courses with verified certificates, while the University of Maryland and the Georgia Institute of Technology each offered two courses with verified certification. Other universities that offered verified certification were the College of St. Scholastica, Duke University, Vanderbilt University, HarvardX, CEU Universidad San Pablo, and Universitat Politecnica de Valencia. Only three courses offered other professional qualifications. Two courses (“Care of Elders with Alzheimer’s Disease and other Major Neurocognitive Disorders” and “Global Tuberculosis (TB) Clinical Management and Research”) offered by John Hopkins University were eligible for Continuing Nursing Education (CNE) Credit while “Caries Management by Risk Assessment (CAMBRA)” by University of California offered 12 units of Continuing Dental Education credit for practicing dental professionals and Continuing Medical Education (CME) Credit for practicing physicians. The summary of certification types offered in MOOC descriptions are shown in Table 7.

Table 7. Certification types on MOOC descriptions.

Type of certification	n	%[a]
Statement of accomplishment	59	91
Statement of accomplishment mentioning no credit awarded	5	8
Verified certificates	14	22
Other professional recognition	3	5

[a] There were 65 MOOCs that offered at least one type of certification.

Prerequisites

Out of the considered courses, 59 course descriptions specifically mentioned whether there were prerequisites or not (Table 8). Some courses used “no prerequisites” or “all are welcome” to describe that the course did not have prerequisites, while some others (eg, “Training and Learning Programs for Volunteer Community Health Workers”) mentioned “Some background in community health programs is helpful but not necessary” (these are categorized under “no prerequisite but helpful background”).

Table 8. Prerequisites in MOOC descriptions (n=59).

Prerequisites	n	%
No prerequisites	17	29
No prerequisites but helpful background	14	24
With prerequisites	28	47

Qualitative Analysis

A word frequency analysis (in NVivo) of course titles (titles in other language were translated using Google Translate) showed that the word most frequently used was “Health” with 35 occurrences (Table 9). The next highest used word was “Introduction” with 13 occurrences. Given that 31 MOOCs had no prerequisites to join, this suggests that many courses offered are introductory level.

Table 9. Frequently used words in MOOC titles.

Word	Frequency
Health	35
Introduction	13
Nutrition	12
Health care	10
Food	8
Nursing	6
Care	5
Clinical	5
Human	5

Target Groups

Only 16 courses specified the target audience for the course. A word frequency analysis (in NVivo) of the audience specified showed that the word most frequently used to define target audience was “health” with 10 occurrences followed by the word “professionals” with 8 occurrences.

Discussion

Health Inequality

This review of MOOCs offered in the area of health and medicine during 2013 provides interesting insights, especially the fact that out of the 98 MOOCs only two were offered by universities in developing countries (“One Health One Medicine” by St. George’s University, Grenada, West Indies, and ‘Traditional Chinese Medicine and Chinese Culture” by Shanghai Jiao Tong University, China). This is not unusual, as to date the large majority of MOOCs are offered by institutes in the developed countries. A contributor for this observation specifically in health and medicine-related MOOCs could be the advanced technologies used in prevention, detection, and treatment in the health care systems of the developed countries and their willingness to showcase the success. On the other hand, it can also be an indicator of health inequality between countries. None of the developing countries’ expertise, for example in tropical diseases, is offered as MOOCs. It is noted that open education resources (OERs) on tropical diseases developed by subject experts in Malawi and Ghana from the African Health OER Network [12] are used in the University of Michigan’s medical programs [13]. Similar collaborations with experts from developing countries/universities on MOOCs may create MOOCs that would be of wider interest. The recent edX partnership with Google to jointly develop the edX open source learning platform perhaps will expand the availability of the platform [14] to individuals and institutions.

Continued Medical Education

Volandes et al [15] argue that online video learning techniques could empower both clinicians and patients. In fact, MOOCs could well be used as a method for Continuing Medical Education (CME). In this review, we found a number of MOOCs that offered verified certification and counted credits toward Continuing Nursing Education, Continuing Dental Education, and CME. Hoy [16] shows that MOOCs can be a convenient and economical method of CME, with the declining industry funding for CME activities.

Medical Student Education

MOOCs can also provide education to students currently undergoing training to become health professionals. For example, the Coursera course “Clinical Terminology for International and US Students” offered by University of Pittsburgh is aimed at new students in the medical field. Courses such as “Going out on a limb: Anatomy of the upper limb” on Coursera by University of Pennsylvania can supplement traditional medical education or perhaps could even be considered as a “flipped-classroom” [17] experience where the MOOC replaces the lecture and the contact hours with the professor used for a more meaningful discussion.

Health Literacy

Health literacy is a broad concept with different definitions. Here, we consider it to be “the degree to which people are able to access, understand, appraise and communicate information to engage with the demands of different health contexts in order to promote and maintain good health across the life-course” [18]. Health literacy, similar to literacy is of critical importance for everyday living [19]. It is not just the ability to make sense of health information but is also a strategy for citizenship and empowerment [19,20]. In this information age, eHealth literacy, or “the ability of people to use emerging information and communications technologies to improve or enable health and health care” [21], is becoming even more relevant.

Specialist information on subjects including health and medicine is becoming widely available today. However, information overload and the availability of unreliable information sources on the Internet present a huge challenge for the general public looking for information on a specific medical condition. Availability of MOOCs (especially if the content is open) is likely to help those who are seeking information. As the content is offered by a reputed institution, it becomes easy to identify it as an authentic and credible source.

Patient Education

MOOCs on health and medicine allow the general public to acquire health education on very specialist topics. One potential area that can be targeted by health and medicine MOOCs is patient education. For example, the MOOC "Care of Elders with Alzheimer's Disease and other Major Neurocognitive Disorders" provides information to anyone who is interested in knowing about Alzheimer's disease. This MOOC welcomes patients in early stages of the disease to help them understand the implications of the disease. Participation in these MOOCs is likely to inform patients of their condition and advanced techniques and interventions that are available. It could, to some extent, bridge the language gap (medical terminology) and knowledge gap between patients and doctors. Thus patients would be able to engage in a meaningful discussion with medical professionals on the care they receive and other alternative treatments.

However, in high power distance cultures, this enlightenment of patients may not be well received by health care professionals. In some instances, informed patients or carers may request treatment not yet available in their contexts. On the other hand, informing patients of possible treatments could facilitate medical tourism for those who can afford it.

Educating the Public

MOOCs can also be used as a tool to educate the public on important health issues. For instance, the Coursera MOOC "Contraception: Choices, Culture and Consequences" offered by University of California educates the public on the importance of reproductive health. Such courses could help people educate themselves without having to talk to a health professional about family planning, which in some cultures is taboo or discouraged by faith. Being able to access trustworthy information through a MOOC could empower people who may otherwise not know the options open to them.

Educating the masses on taboo topics such as "Drug Addiction" could also be achieved with the use of MOOCs. MOOCs generally operate entirely on online spaces; exceptions would be MOOC participants from a locality meeting up for discussions or MOOC participants seeking support from blended provision such as Coursera Learning Hubs [22]. Because one can project a persona in an online space that may differ to who they are in real life, both patients with such conditions and others who seek information can join in, if they wish, without revealing their true identity.

However, suggesting that MOOCs may be a way to educate the public assumes that other conditions for participation in a MOOC (eg, access to technology, skills to use them, and international language competency) are already met. But as Liyanagunawardena, Williams, and Adams [23] show, at present MOOCs may not reach a large proportion of people, especially in developing areas of the world. Current data suggest that a typical course registrant "is a male with a bachelor's degree who is 26 or older" [24], showing that MOOCs have not yet reached universal accessibility.

Limitations

This review was conducted by collecting data from various sources. However, as very few MOOC platforms provided official data on their MOOC offerings, only the courses with publicly available course details were used in the review. Collection of data for the review using aggregator sites could have the disadvantage of not including all MOOCs on offer, especially foreign language offerings. By using two aggregator sites and independently collecting data from MOOC platforms, the possibility of this occurring was minimized. In collecting MOOCs for the review, courses categorized under health and medicine or related was considered. However, if a MOOC were wrongly classified, it would not have been captured in the review. In instances where the MOOCs were offered in languages other than English, Google Translate was used to translate the content. Had there been a translation error, it could have affected the data collection process.

Conclusions

Massive open online courses (MOOCs) have become popular within a short span of time, and there are dozens of providers offering courses in a variety of subjects. Reviewing MOOCs offered on "Health and Medicine" in 2013, we found that 94% of them (92/98) were offered in English and the large majority were offered by North American institutions. Only about 3% of the MOOCs (3/98) were offered by institutions in the developing world. Many courses offered were introductory level. Some courses offered credit toward continuous professional development of medical professionals and verified certificates for a fee, while others offered a statement of participation for successful participants.

There is potential to use MOOCs to educate health care practitioners and students; for example, in continuous professional development of health care professionals. Because they can reach massive numbers across the globe, MOOCs can provide an enormous boost in educating the public on health and medicine, especially on taboo subjects such as acquired immunodeficiency syndrome (AIDS), tuberculosis, and contraception. However, in order to unleash this great potential of MOOCs in educating masses around the globe on health and wellbeing, there are various challenges to overcome (eg, access: language access, physical access to technology, skills access to use technology). Health literacy is a powerful tool that empowers people, and MOOCs could be used to educate the general public to increase their health literacy. The wide variety of MOOCs on various subjects relating to health and medicine offered in 2013 show a glimpse of what is achievable through MOOCs in this discipline.

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful comments on an earlier version of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MOOC platforms.

[PDF File (Adobe PDF File), 4KB - jmir_v16i8e191_app1.pdf]

Multimedia Appendix 2

List of MOOCs considered in review.

[PDF File (Adobe PDF File), 27KB - jmir_v16i8e191_app2.pdf]

References

1. Downes S. Innovate: Journal of Online Education. 2008. Places to go: Connectivism & connective knowledge URL: http://bsili.3csn.org/files/2010/06/Places_to_Go-__Connectivism__Connective_Knowledge.pdf [accessed 2014-08-09] [WebCite Cache ID 6Rh66Z5aH]
2. Liyanagunawardena TR, Adams AA, Williams SA. MOOCs: a Systematic Study of the Published Literature 2008-2012. International Review of Research in Open and Distance Learning 2013;14(3):201-227 [FREE Full text]
3. Siemens G. Connectivism - A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning 2005;2(1):3-10.
4. Yuan L, Powell S. JISC ECTIS (Centre for Educational Technology and Interoperability Standards). 2013. MOOCs and Open Education: Implications for Higher Education – A white paper URL: http://publications.cetis.ac.uk/wp-content/uploads/2013/03/MOOCs-and-Open-Education.pdf [accessed 2014-07-15] [WebCite Cache ID 6R5L67mhj]
5. Liyanagunawardena TR, Parslow P, Williams SA. Dropout: MOOC participants'perspective. In: Proceedings of EMOOCs 2014, the Second MOOC European Stakeholders Summit. 2014 Presented at: EMOOCs 2014 the Second MOOC European Stakeholders Summit; February 10-12, 2014; Lausanne, Switzerland p. 95-100.
6. Downes S. What the 'x' in 'xMOOC' stands for. 2013 Apr 09. URL: https://plus.google.com/109526159908242471749/posts/LEwaKxL2MaM [accessed 2014-03-19] [WebCite Cache ID 6OC7GFwKK]
7. Rodriguez CO. MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. European Journal of Open, Distance and E-Learning 2012;2012(II):1-13 [FREE Full text]
8. Mackness J. Blog. 2013 Oct 11. cMOOCs and xMOOCs – key differences URL: http://jennymackness.wordpress.com/2013/10/22/cmoocs-and-xmoocs-key-differences/ [accessed 2014-03-19] [WebCite Cache ID 6OC8OPiOo]
9. List Education Technology. 2013 Aug 19. Worldwide MOOCs (19 systems) URL: http://listedtech.com/content/worldwide-moocs-19-systems [accessed 2014-03-19] [WebCite Cache ID 6OC8lzmwx]
10. Class Central. URL: https://www.class-central.com/ [accessed 2014-03-20] [WebCite Cache ID 6ODGxIEhi]
11. MOOC List. URL: http://www.mooc-list.com/ [accessed 2014-03-20] [WebCite Cache ID 6ODH92L45]
12. African Health OER Network. URL: http://www.oerafrica.org/healthoer [accessed 2014-03-20] [WebCite Cache ID 6ODMx43yc]
13. Vollmer T. Creative Commons. Sir John Daniel of the Commonwealth of Learning: Open Education and Policy URL: http://creativecommons.org/weblog/entry/28384 [accessed 2014-03-20] [WebCite Cache ID 6ODNA5KuW]
14. mooc.org. URL: http://mooc.org/ [accessed 2014-03-20] [WebCite Cache ID 6ODNLUqRa]
15. Volandes AE, Kennedy WJ, Davis AD, Gillick MR, Paasche-Orlow MK. The new tools: What 21st century education can teach us. Healthcare 2013 Dec;1(3-4):79-81. [doi: 10.1016/j.hjdsi.2013.07.011]
16. Hoy MB. MOOCs 101: an introduction to massive open online courses. Med Ref Serv Q 2014 Jan;33(1):85-91. [doi: 10.1080/02763869.2014.866490] [Medline: 24528267]
17. Tucker B. Education Next. 2012. The Flipped Classroom: Online instruction at home frees class time for learning URL: http://educationnext.org/files/ednext_20121_BTucker.pdf [accessed 2014-08-09] [WebCite Cache ID 6RhBj6oBZ]
18. Kwan B, Frankish J, Rootman I. University of British Columbia. 2006 Nov. The development and validation of measures of "health literacy" in different populations. Institute of Health Promotion Research & University of Victoria Centre for Community Health Promotion Research URL: http://blogs.ubc.ca/frankish/files/2010/12/HLit-final-report-2006-11-24.pdf [accessed 2014-03-20] [WebCite Cache ID 6ODPMpsyk]
19. Kanj M, Mitic W. Working Papers - 7th Global Conference on Health Promotion. Individual Empowerment URL: http://www.who.int/healthpromotion/conferences/7gchp/Track1_Inner.pdf [accessed 2014-03-20] [WebCite Cache ID 6ODRFi4F5]
20. Patient Information Forum. Health literacy and health information producers URL: http://goo.gl/hTy1f9 [accessed 2014-03-19] [WebCite Cache ID 6ODSTtWtP]
21. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]

22. Coursera. Coursera Learning Hubs URL: https://www.coursera.org/about/programs/learningHubs [accessed 2014-07-16] [WebCite Cache ID 6R6YjNUlY]
23. Liyanagunawardena T, Williams S, Adams A. The Impact and Reach of MOOCs: A Developing Countries' Perspective. eLearning Papers (MOOCs and Beyond) 2013;33:1-8 [FREE Full text]
24. Ho AD, Reich J, Nesterko S, Seaton DT, Mullaney T, Waldo J, et al. HarvardX and MITx Working Paper No. 1. 2014. HarvardX and MITx: The first year of open online courses URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2381263 [accessed 2014-08-09] [WebCite Cache ID 6RhCuclWM]

Abbreviations

CME: Continued Medical Education
cMOOC: connectivist MOOC
MOOC: massive open online course
OER: Open Educational Resources
xMOOC: MOOC as eXtension of something else

Edited by G Eysenbach; submitted 02.04.14; peer-reviewed by M Hoy, K Murphy; comments to author 14.07.14; revised version received 23.07.14; accepted 25.07.14; published 14.08.14.

Please cite as:
Liyanagunawardena TR, Williams SA
Massive Open Online Courses on Health and Medicine: Review
J Med Internet Res 2014;16(8):e191
URL: http://www.jmir.org/2014/8/e191/
doi:10.2196/jmir.3439
PMID:25123952

Original Paper

Physiological and Brain Activity After a Combined Cognitive Behavioral Treatment Plus Video Game Therapy for Emotional Regulation in Bulimia Nervosa: A Case Report

Ana Beatriz Fagundo[1,2], PhD; Esther Via[1], MD; Isabel Sánchez[1], MSc; Susana Jiménez-Murcia[1,2,3], PhD; Laura Forcano[1,2], PhD; Carles Soriano-Mas[1,4], PhD; Cristina Giner-Bartolomé[1], MSc; Juan J Santamaría[1,2], MSc; Maher Ben-Moussa[5], MSc; Dimitri Konstantas[5], PhD; Tony Lam[6], MSc; Mikkel Lucas[6], MSc; Jeppe Nielsen[7], PhD; Peter Lems[8], MSc; Narcís Cardoner[1,3,4], PhD; Jose M Menchón[1,3,4], PhD; Rafael de la Torre[2,9], PhD; Fernando Fernandez-Aranda[1,2,3], PhD

[1]University Hospital of Bellvitge-IDIBELL, Department of Psychiatry, Barcelona, Spain

[2]Instituto Salud Carlos III, CIBER Fisiopatología Obesidad y Nutrición (CIBERObn), Barcelona, Spain

[3]School of Medicine, Department of Clinical Sciences, University of Barcelona, Barcelona, Spain

[4]Instituto Salud Carlos III, CIBER Salud Mental (CIBERSam), Barcelona, Spain

[5]University of Geneva, Geneva, Switzerland

[6]NetUnion, Laussane, Switzerland

[7]Serious Game Interactive (SGI), Copenhagen, Denmark

[8]MobiHealth BV, Enschede, Netherlands

[9]IMIM (Hospital del Mar Medical Research Institute), Human Pharmacology and Clinical Neurosciences Research Group, Neuroscience Research Program, Barcelona, Spain

Corresponding Author:
Fernando Fernandez-Aranda, PhD
University Hospital of Bellvitge-IDIBELL
Department of Psychiatry
Feixa Llarga s/n
Barcelona, 08907
Spain
Phone: 34 93 2607227
Fax: 34 93 2607193
Email: ffernandez@bellvitgehospital.cat

Abstract

Background: PlayMancer is a video game designed to increase emotional regulation and reduce general impulsive behaviors, by training to decrease arousal and improve decision-making and planning. We have previously demonstrated the usefulness of PlayMancer in reducing impulsivity and improving emotional regulation in bulimia nervosa (BN) patients. However, whether these improvements are actually translated into brain changes remains unclear.

Objective: The aim of this case study was to report on a 28-year-old Spanish woman with BN, and to examine changes in physiological variables and brain activity after a combined treatment of video game therapy (VGT) and cognitive behavioral therapy (CBT).

Methods: Ten VGT sessions were carried out on a weekly basis. Anxiety, physiological, and impulsivity measurements were recorded. The patient was scanned in a 1.5-T magnetic resonance scanner, prior to and after the 10-week VGT/CBT combined treatment, using two paradigms: (1) an emotional face-matching task, and (2) a multi-source interference task (MSIT).

Results: Upon completing the treatment, a decrease in average heart rate was observed. The functional magnetic resonance imaging (fMRI) results indicated a post-treatment reduction in reaction time along with high accuracy. The patient engaged areas typically active in healthy controls, although the cluster extension of the active areas decreased after the combined treatment.

Conclusions: These results suggest a global improvement in emotional regulation and impulsivity control after the VGT therapy in BN, demonstrated by both physiological and neural changes. These promising results suggest that a combined treatment of CBT and VGT might lead to functional cerebral changes that ultimately translate into better cognitive and emotional performances.

(J Med Internet Res 2014;16(8):e183) doi:10.2196/jmir.3243

KEYWORDS

eating disorders; bulimia nervosa; emotional regulation; impulsivity; video game therapy; neuroimaging; fMRI

Introduction

Conventional psychological therapies, such as cognitive behavioral therapy (CBT) have been successful in treating some central symptoms in eating disorders, such as binge/purging behaviors [1]. However, in bulimia nervosa (BN) patients, there are some dysfunctional features that still remain after treatment. Some of them, such as alterations in executive functioning (ie, impulsivity, planning, and decision making) and emotional deregulation (eg, self-control strategies or tolerance to frustration) are particularly difficult to modify and are associated with an adverse outcome [2].

PlayMancer is a video game designed to increase emotional regulation and reduce general impulsive behaviors, by training to decrease arousal and improve decision-making and planning [3,4]. As described in previous research [5], the final aim of video game therapy (VGT) is to achieve more efficient brain functioning with appropriate emotional and cognitive processing, which eventually translates into more suitable real-world behaviors. We have previously demonstrated the usefulness of PlayMancer in reducing impulsivity and enhancing emotional regulation in BN patients [6]. However, whether these behavioral improvements are also translated into changes in brain activity remains unclear.

This case report aimed at examining changes in physiological reactivity and brain activation as biomarkers of emotional regulation and impulsivity control in response to a combination of VGT and CBT treatments in a BN patient.

Methods

Case Report

Overview

The research procedures were explained in full to the patient and she gave written informed consent prior to enrollment in the study. The procedures were approved by the Ethical Committee of the University Hospital of Bellvitge. The patient was a 28-year-old woman seeking treatment for BN in our outpatient unit. BN was diagnosed according to the DSM-5 criteria [7]. The BN disorder started at the age of 22 (age of onset), after the patient had followed a hypocaloric diet. Pre-morbid overweight, a body mass index (BMI) of 28.7, and various psychosocial stressors were starter risk factors. The patient reported having started BN with 1-2 weekly bingeing and vomiting episodes. During approximately the last 6 months prior to the initiation of the treatments, the patient exhibited more than four weekly binge-eating episodes with compensatory fasting/restricting eating behaviors but without purging episodes, and continued to have extreme concerns about shape and weight. At the time of inclusion in the case study, her weight was 69.3 kg (height 166 cm, BMI 25.1).

Personal and Psychiatric Antecedents

The patient, the younger of two children, is currently living with her partner. Moreover, at the start of the treatment she did not present further additional psychiatric comorbidity or other familial psychiatric disorders, alcohol-drug misuse, regular tobacco consumption, or any other relevant difficulties for dealing with stress and negative emotions.

Psychometric Assessment and Physiological Measures

At the beginning, the patient was given the Eating Disorder Inventory [8], Symptom Checklist-90-Revised (SCL-90-R) [9], and Temperament and Character Inventory-Revised [10]. The psychometric assessment revealed a typical profile described frequently by BN patients (characterized by high body dissatisfaction, drive for thinness, bulimic episodes, social insecurity, anxiety, mild depressive symptoms, high harm avoidance, and low self-directedness). Comorbidity was assessed by means of the structured clinical interview for DSM-IV Axis I disorders (SCID-I/II). Additionally, weekly binge-eating and purging frequencies were recorded and monitored by means of a food diary throughout the duration of the therapy. Prior to and after VGT treatment, State-Trait Anxiety Inventory (STAI) [11] and Barratt Impulsiveness (BIS-11) [12] scales were administered. The physiological measures were analyzed by means of a sensor system via Bluetooth, linked to PlayMancer, including among others, autonomic measurements such as heart rate, pulse rate, and heart rate variability measures.

Functional Magnetic Resonance Imaging Procedure

Paradigms

Emotional Face-Matching Task

To explore emotional activation and regulation, we used a modified version of the emotional face-matching task originally reported by Hariri, Bookheimer, and Mazziotta [13], which has also been reported elsewhere [14]. This task was proved to reliably activate the visual cortex, the amygdala, and the dorsolateral prefrontal cortex in healthy subjects. The contrasts of interest were fearful faces to shapes and happy faces to shapes. During each 5-second trial, the patient was presented with a target face (center top) and two probe faces (bottom left and right) and was instructed to match the probe expressing the same emotion to the target by pressing a button in either their left or right hand of a magnetic resonance imaging (MRI) compatible response device. The target face was either happy or fearful, and the probe faces included two out of three possible emotional faces (happy, fearful, and angry). As a sensorimotor control condition, the patient was presented with 5-second trials of ovals or circles in an analogous configuration and was instructed to match the shape of the probe to the target. A total of six 30-second blocks of faces (3 fearful, 3 happy) and six 30-second blocks of the control condition (shapes) were presented interleaved in a pseudo-randomized order. The contrasts of interest were fearful faces to shapes and happy faces to shapes.

Executive-Control Task

As an executive-control task, we employed the multi-source interference task (MSIT) [15], a task that reliably and robustly activates cingulo-frontal-parietal cognitive/attention network. Our contrast of interest was the incongruent condition to congruent condition. During each trial, the patient was asked to press one of three buttons in an MRI-compatible response device to identify the unique digit (1, 2, or 3) that was not repeated in a string of three digits. Each digit was mapped to index (1), middle (2), and ring (3) fingers of the right hand and subjects should respond to the identity (not the position) of the unique digit in the string. There were a total of 48 congruent trials (ie, 133, 121, 223) and 48 incongruent trials (ie, 212, 311, 322), distributed in a total of eight blocks of 12 trials each, separated with nine fixation crosses of 15-second duration. In the incongruent situation, the unique digit's spatial position in the string was conflicted with the corresponding identity of the unique digit. Our contrast of interest was the incongruent condition to congruent condition.

Acquisition, Processing, and Analyses of the Images

The patient was scanned twice in a 1.5-T Signa Excite system (General Electric) Magnetic Resonance (MR) scanner, prior to and after the 10-week VGT/CBT combined treatment. The MR was equipped with an 8-channel phased-array head coil and single-shot echoplanar imaging software was used. The functional sequence consisted of gradient recalled acquisition in the steady state (repetition time=2000 ms, echo time=50 ms and pulse angle, 90º) in a 24 cm field of view, with a 64 x 64 pixel matrix, and a slice thickness of 4 mm (inter-slice gap, 1.5 mm).

A total of 22 interleaved sections, parallel to the anterior-posterior commissure line, were acquired to generate 207 (MSIT task) and 195 (emotional face-matching task) whole-brain volumes. Visual stimuli were presented using MRI-compatible goggles (VisuaStim Digital System, Resonance Technology Inc, Northridge, CA, USA), while behavioral responses were recorded by means of an MRI-compatible response grip (NordicNeuroLab Inc, Bergen, Norway). Imaging data were processed on a Macintosh platform running Matlab version 7 (The MathWorks Inc) and statistical parametric mapping software version 8 (SPM8). Time series of the pre and post acquired images were initially realigned to the mean image by using a 6-parameter (rigid body) spatial transformation, normalized to the standard eco-planar imaging template in SPM, resliced in Montreal Neurological Institute (MNI) space and smoothed using a Gaussian filter (full width at half maximum=8 mm). Realigned, normalized, and smoothed images were carried to a first level of analysis for the contrasts of interest in each task. Results were thresholded at a positive false discovery rate<0.05 corrected. Finally, the pre and post VGT/CBT combined treatment were overlapped in a T1 MNI template in MRIcron software (Rorden and Brett, 2000) for visual comparison and presentation of the results.

Treatment

As described previously [6], a combined therapy was used in this case study (16 weekly outpatient cognitive-behavioral group sessions plus 10 weekly sessions of VGT) (Figure 1). A detailed description of the main goal, techniques, and structure of the CBT group therapy have been described previously [1], as well as the VGT approach used [3,4].

The performance in each VGT session was collected during 20 minutes. Three minutes of relaxing music were played before and after the VGT session. The video game consisted of three mini-games: (1) The Face of Cronos, where the player has to climb up a cliff in which obstacles appear depending on the arousal of the player (based on biofeedback); this mini-game trains planning and decision making, (2) Treasures of the Sea, which is a virtual swimming game in which the player has to collect different objects and fishes while conserving their oxygen supply; this trains visuospatial abilities, visual working memory, and decision making, and high arousal makes the task more difficult, and (3) Sign of the Magupta, which is a relaxation game in which the player connects a constellation of stars through breathing control; slow deep breathing allows the connections between stars to form [6] (see Figure 2 and Multimedia Appendix 1).

Figure 1. Schematic diagram of treatments and procedures.

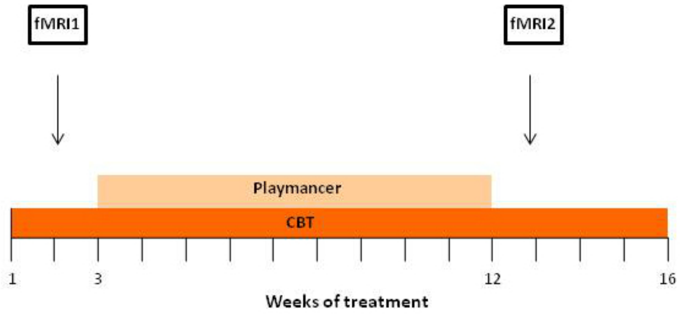
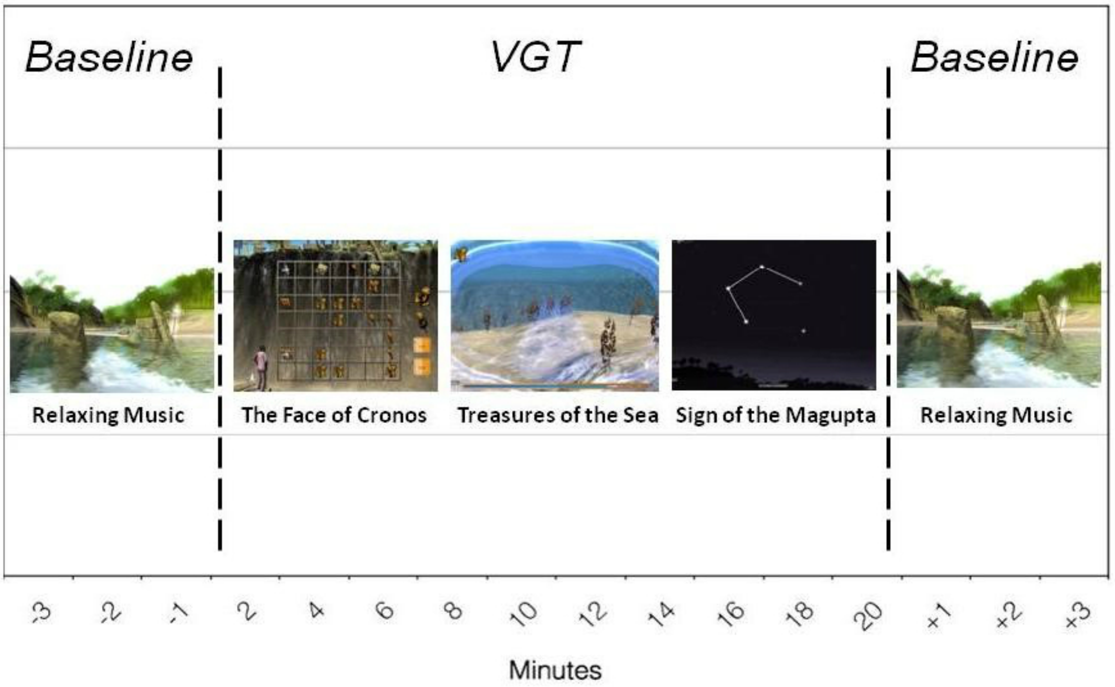


Figure 2. Example of PlayMancer session: relaxing phases and mini-games.



Results

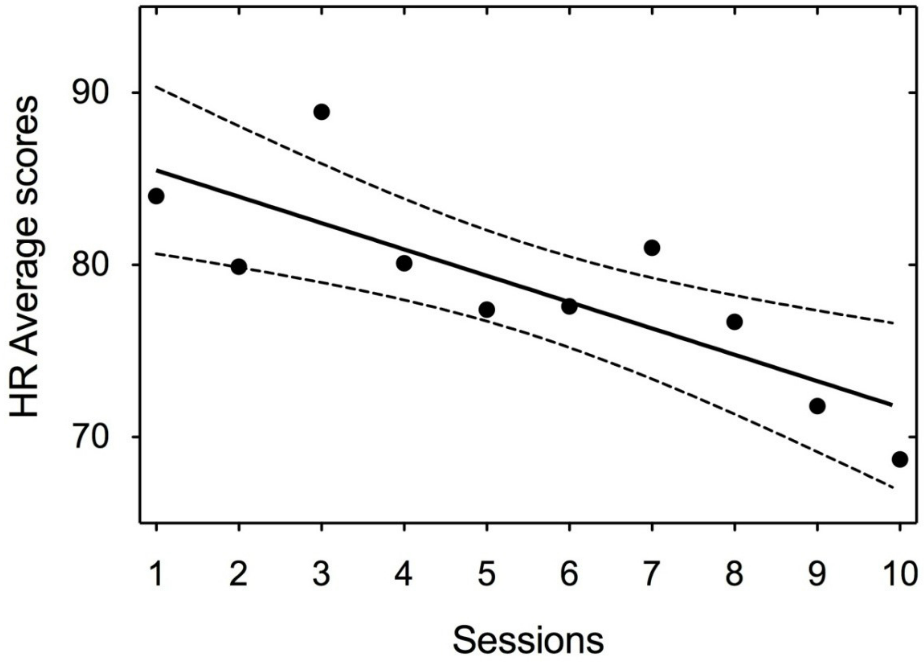
Eating Symptomatology, Anxiety, and Impulsiveness

During the combined treatment, the rate of binge eating and the consequent compensatory behaviors started to decline after the fourth session, whereas abstinence of bingeing occurred after the sixth session. At the 6- and 12-month follow-ups, the patient was still free of binge eating symptoms. Regarding secondary outcome measures (namely anxiety and impulsivity), the patient had reduced impulsivity (measured by means of BIS-11, pre mean score 38 and post mean score 29, with the Spanish population mean score of 32.5, Oquendo et al, 2001) and state of anxiety (measured by means of STAI, pre mean 38, post mean 17) after VGT intervention, and it was maintained after follow-up.

Evolution of Physiological Variables With the Video Game Therapy

A trend was found for the weekly average heart rate (HR), which signified scores tended to decrease with the game sessions (Figure 3).

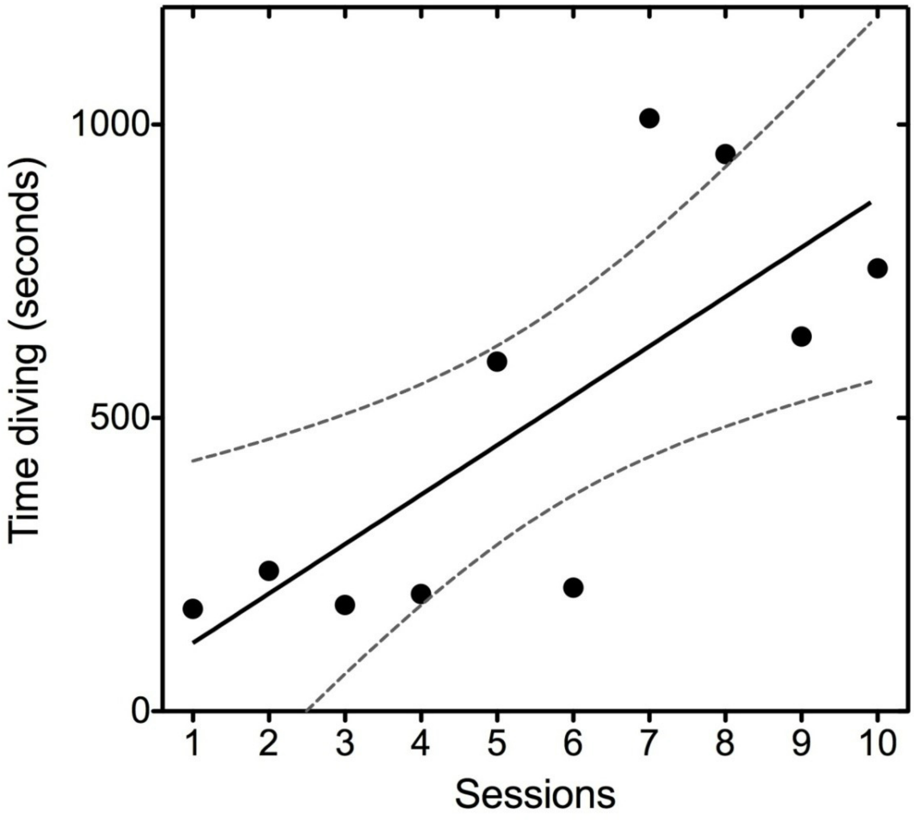
Figure 3. Average Heart Rate (HR) scores over the 10 therapy sessions.



Internal Outcome Measure of the Video Game Therapy

One of the internal VGT measures was the total time spent in a specific task (diving task), which is the absolute diving time (in seconds) the subject plays without being interrupted (due to lack of stress management or emotional regulation capacity, oxygen ran out, and the diving session was interrupted). A positive trend was found for the weekly average total diving time, which means scores tended to increase with the sessions of the game (Figure 4).

Figure 4. Video game internal secondary outcome measure (average time diving in seconds over the 10 therapy sessions).



Pre-Post Changes in Brain Activity

Behavioral measures indicated that reaction time decreased in the post condition in both the face-matching task and the MSIT task (see Table 1). The patient had high global accuracy in both tasks (face-matching tasks: pre: 97.22%, post: 98.61%; MSIT: pre: 100%, post: 97.92%).

The face-matching task engaged areas typically active during this task in healthy controls, such as fusiform gyrus, visual cortex, precentral and dorsolateral prefrontal cortices (see Figure 5). Pre and post conditions were partially overlapping, although, globally, the cluster extension of the active areas decreased after combined treatment. A small cluster in the left amygdala (2 voxels, see Figure 5) was also active during the fearful faces matching, even though only in the pre treatment condition (matching happy faces during the pre condition and fearful and happy faces during the post condition did not activate the amygdala, even when lowering the threshold to an uncorrected *P*<.001). Other differences between pre and post conditions were located in the bilateral dorsolateral prefrontal cortex during the matching of fearful faces and in the bilateral frontopolar-anterior insula /bilteral dorsolateral prefrontal cortex during the matching of happy faces.

With regard to the MSIT task, the incongruent versus congruent conditions showed activations in the medial prefrontal-dorsal anterior cingulate and superior parietal cortex, for both pre and post assessments (see Figure 5). Although these areas are typically activated in healthy controls while carrying out the task, again at post assessment the patient engaged a smaller cluster extend of voxels. Additional activations were found in the bilateral anterior insula, which were also found to decrease after the combined treatment (Figure 6).

Table 1. Pre-post changes in brain activity (fMRI[a]): behavioral measures.

Paradigm		Pre, mean (SD)[b]	Post, mean (SD)[b]
Face-matching task			
	Fearful faces trials	1.44 (0.07)	1.12 (0.04)
	Happy faces trials	1.00 (0.05)	0.91 (0.03)
	Shapes trials	0.72 (0.03)	0.68 (0.08)
Multi source interference task (MSIT)			
	Congruent trials	0.542 (0.009)	0.478 (0.005)
	Incongruent trials	1.104 (0.063)	0.917 (0.082)

[a] fMRI: functional magnetic resonance imaging.

[b] Results are presented in seconds.

Figure 5. Pre-treatment and post-treatment activations during the emotional face-matching task (happy faces vs shapes and fearful faces vs shapes) and the multi-source interference (MSIT) task (incongruent vs congruent condition).

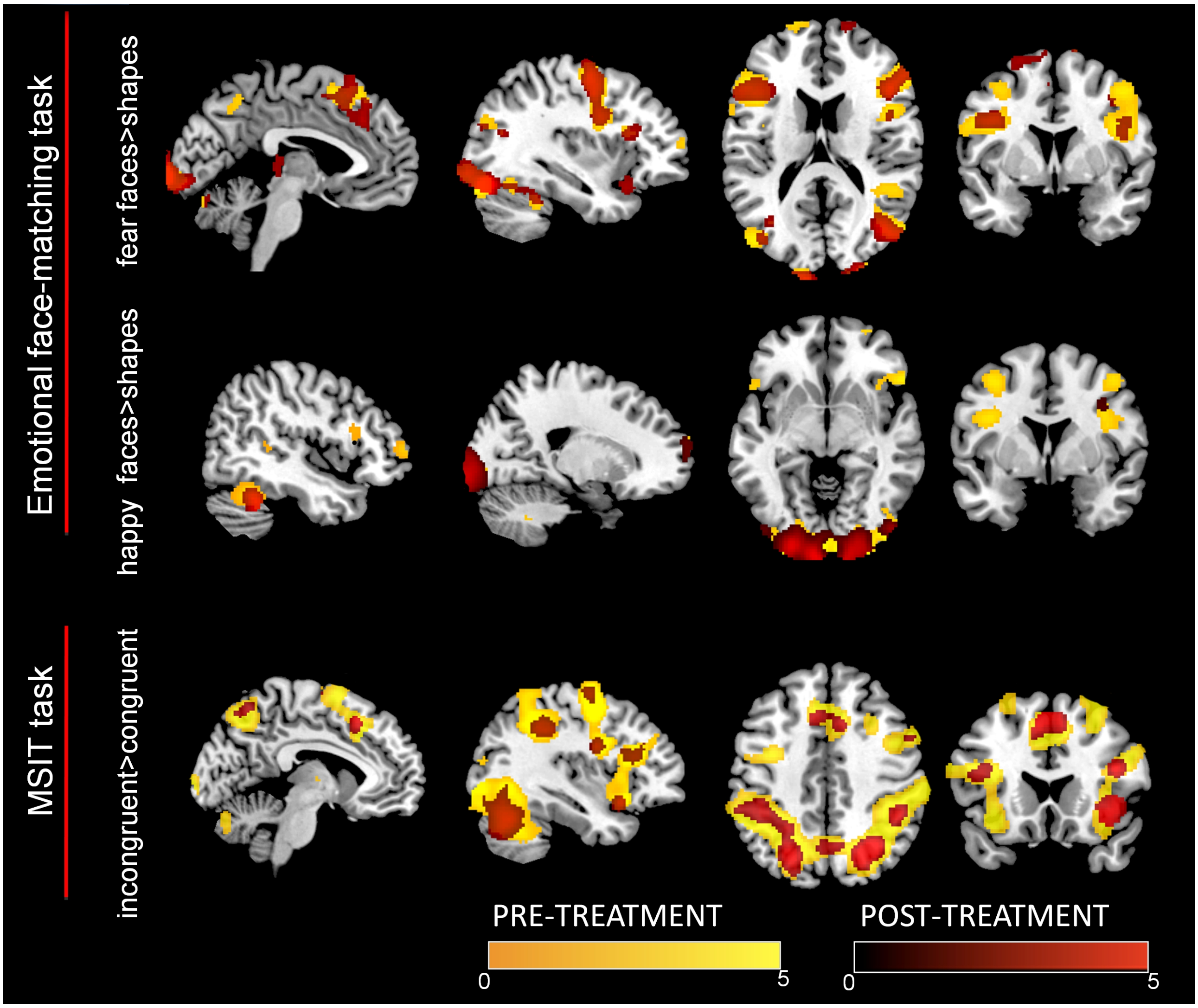
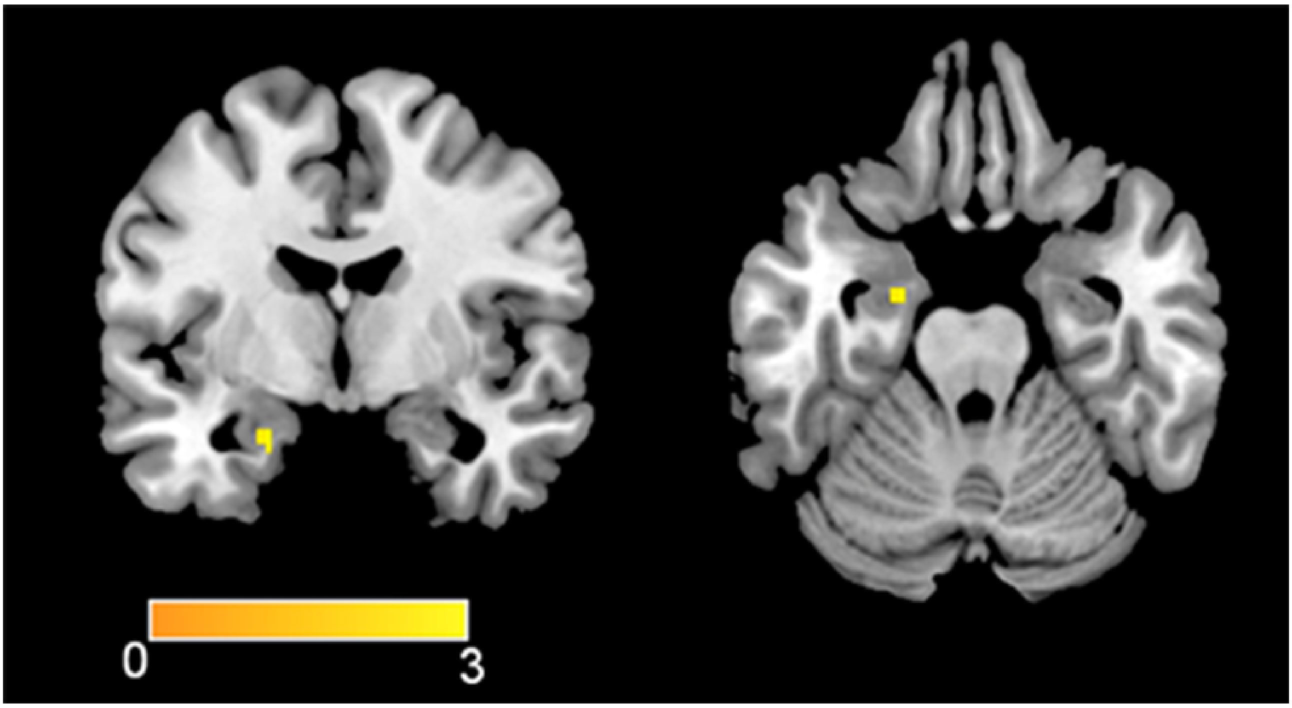


Figure 6. Activations in the bilateral anterior insula after the treatment.



Discussion

Principal Findings

This case study set out to examine the neural and physiological changes associated with a combined therapy (CBT plus video game therapy), as a tool to enhance emotional regulation and impulsivity control in a BN patient. In this reported case, the results suggest that specific training to decrease arousal and increase impulsivity control [6] may improve emotional regulation, and induce changes in physiological variables (eg, HR) and in the neural circuits related to emotional and executive processing. Although emotional regulation has been previously studied in BN [6], and new technologies have been previously used in psychology [16,17], this is, to the best of our knowledge, the first time that the neural and physiological changes associated with a combination of VGT and CBT have been described in a BN patient.

According to our results, self-regulation training incorporating physiological and emotional feedback might improve the emotional regulation capacity. Specifically, the reduction of physiological markers (eg, HR) after the VGT may suggest higher emotional control and is in agreement with those studies showing that a dysfunctional control over emotions is associated with increased heart rate, heart rhythm dysregulation, and autonomic imbalance [18,19]. In this regard, recent evidence indicates that decreases in physiological variables, such as HR, are connected not only to a higher self-regulation of emotions, but also with specific improvements in executive functions (eg, cognitive flexibility and control over impulsive behaviors) [20], which is also in line with the neuroimaging results in this case report.

In this sense, after the combined VGT/CBT treatment, the patient was able to display a better behavioral performance on both the video game tasks and the functional magnetic resonance imaging (fMRI) paradigms. Specifically, an enhancement of absolute time playing the game and lower reaction times with high global accuracy in the fMRI paradigm were observed. As explained above, PlayMancer is a video game specifically designed for training emotional regulation, but can also be used for training planning, inhibition response, decision making, and working memory capacities [6]. Thus, these findings suggest that in this BN case, the combination of CBT with regular cognitive training improves cognitive functions and produces changes in the neural substrates associated.

Specifically, the neural patron displayed by the patient while performing the fMRI paradigms was similar to the one found in healthy controls, for both the impulsive control paradigm (MSIT) [15] and the emotional paradigm [14], but was modified after the combined treatment. Executive and attention networks were active during emotion processing and cognitive interference tasks. However, on the one hand, it is suggested that brain activations were more efficient after the combined treatment, given that an improvement of behavioral results was achieved with a lesser extent of cluster activations. The amygdala also showed some differences between the pre and post conditions of the fearful face-matching trials. Although the cluster extent of the amygdalar activation was small, it may also suggest some differences in emotional regulation between the pre and post treatment conditions.

On the other hand, the activation of the anterior insula during cognitive conflict (incongruent condition versus congruent condition) may be consistent with the idea that additional brain resources were needed to perform the tasks. Even though the anterior insula is not generally activated by the MSIT [15], it has been found to be active during performance monitoring and is modulated by error awareness [15]. Additionally, the anterior insula is suggested to be a key dysfunctional structure in the pathophysiology of eating disorders [21].

In summary, these promising results suggest that, in this case report, a combined treatment of CBT and VGT might lead to functional cerebral changes that might eventually translate into better cognitive and emotional performances. This report emphasizes the importance of researching new treatments for enhancing emotional regulation and impulsivity control in BN patients.

Strengths and Limitations

This case study also has several important strengths, primarily the novelty of the therapeutic approach. VGT, as applied in the present case study, might be a practical tool for the treatment of cognitive and emotional alterations in BN.

However, the results of this case study should be interpreted within the context of some limitations. The most important one is that it is a single case report study, although a longitudinal design was employed and pre-post measures of the patient were considered. However, brain activity pre-post was not statistically compared, thus future studies evaluating series of cases-controls should be conducted in order to confirm these findings. Additionally, although the repetition of the task may be contributing to the improvement in brain and behavioral efficiency, this effect could not be differentiated here and would need to be further tested in a case-control study.

Conclusions

Though this report exemplifies a novel treatment for cognitive and emotional rehabilitation in BN patients (ie, video game therapy), more studies need to be carried out and future neuropsychological and neuroimaging studies should focus on the executive and emotional profile of these patients, in order to shed more light on these multifaceted constructs.

Acknowledgments

Financial support was received from Fondo de Investigación Sanitaria-FIS (PI081714; PI11/210); AGAUR (2009SGR1554, 2009SGR718), and Ministerio de Ciencia e Innovación Subprograma Juan de la Cierva (JCI-2011-09248). CIBER Fisiopatología de la Obesidad y Nutrición (CIBERobn) and CIBER Salud Mental (CIBERSAM) are both an initiative of ISCIII. Dr Soriano-Mas is supported by a Miguel Servet contract from the Carlos III Health Institute (CP10/00604).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Video game PlayMancer.

[MP4 File (MP4 Video), 17MB - jmir_v16i8e183_app1.mp4]

References

1. Agüera Z, Riesco N, Jiménez-Murcia S, Islam MA, Granero R, Vicente E, et al. Cognitive behaviour therapy response and dropout rate across purging and nonpurging bulimia nervosa and binge eating disorder: DSM-5 implications. BMC Psychiatry 2013;13:285 [FREE Full text] [doi: 10.1186/1471-244X-13-285] [Medline: 24200085]
2. Dahlgren CL, Lask B, Landrø NI, Rø Ø. Patient and parental self-reports of executive functioning in a sample of young female adolescents with anorexia nervosa before and after cognitive remediation therapy. Eur Eat Disord Rev 2014 Jan;22(1):45-52. [doi: 10.1002/erv.2265] [Medline: 24129969]
3. Fernández-Aranda F, Jiménez-Murcia S, Santamaría JJ, Gunnard K, Soto A, Kalapanidas E, et al. Video games as a complementary therapy tool in mental disorders: PlayMancer, a European multicentre study. J Ment Health 2012 Aug;21(4):364-374 [FREE Full text] [doi: 10.3109/09638237.2012.664302] [Medline: 22548300]
4. Jiménez-Murcia S, Fernández-Aranda F, Kalapanidas E, Konstantas D, Ganchev T, Kocsis O, et al. Playmancer project: a serious videogame as an additional therapy tool for eating and impulse control disorders. Stud Health Technol Inform 2009;144:163-166. [Medline: 19592756]
5. Claes L, Jiménez-Murcia S, Santamaría JJ, Moussa MB, Sánchez I, Forcano L, et al. The facial and subjective emotional reaction in response to a video game designed to train emotional regulation (Playmancer). Eur Eat Disord Rev 2012 Nov;20(6):484-489. [doi: 10.1002/erv.2212] [Medline: 23097370]
6. Fagundo AB, Santamaría JJ, Forcano L, Giner-Bartolomé C, Jiménez-Murcia S, Sánchez I, et al. Video game therapy for emotional regulation and impulsivity control in a series of treated cases with bulimia nervosa. Eur Eat Disord Rev 2013 Nov;21(6):493-499. [doi: 10.1002/erv.2259] [Medline: 24092598]
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th edition. Arlington, VA: American Psychiatric Publishing; 2013.
8. Garner DM. Eating Disorder Inventory-2. Odessa: Psychological Assessment Resources; 1991.
9. Derogatis LR. SCL-90-R. Administration, scoring and procedures manual. Baltimore, MD: Clinical Psychometric Research; 1990.

10. Cloninger CR. The Temperament and Character Inventory–Revised. St Louis, MO: Center for Psychobiology of Personality, Washington University; 1999.
11. Spielberger CD, Gorsuch RL, Lushene RE. Cuestionario de Ansiedad Estado Rasgo, STAI. Madrid: TEA Ediciones; 1982.
12. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. J Clin Psychol 1995 Nov;51(6):768-774. [Medline: 8778124]
13. Hariri AR, Bookheimer SY, Mazziotta JC. Modulating emotional responses: effects of a neocortical network on the limbic system. Neuroreport 2000 Jan 17;11(1):43-48. [Medline: 10683827]
14. Pujol J, Harrison BJ, Ortiz H, Deus J, Soriano-Mas C, López-Solà M, et al. Influence of the fusiform gyrus on amygdala response to emotional faces in the non-clinical range of social anxiety. Psychol Med 2009 Jul;39(7):1177-1187. [doi: 10.1017/S003329170800500X] [Medline: 19154647]
15. Bush G, Shin LM. The Multi-Source Interference Task: an fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network. Nat Protoc 2006;1(1):308-313. [doi: 10.1038/nprot.2006.48] [Medline: 17406250]
16. Gorini A, Gaggioli A, Vigna C, Riva G. A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008;10(3):e21 [FREE Full text] [doi: 10.2196/jmir.1029] [Medline: 18678557]
17. Stubbings DR, Rees CS, Roberts LD, Kane RT. Comparing in-person to videoconference-based cognitive behavioral therapy for mood and anxiety disorders: randomized controlled trial. J Med Internet Res 2013;15(11):e258 [FREE Full text] [doi: 10.2196/jmir.2564] [Medline: 24252663]
18. Gerber AJ, Posner J, Gorman D, Colibazzi T, Yu S, Wang Z, et al. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. Neuropsychologia 2008;46(8):2129-2139 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2008.02.032] [Medline: 18440572]
19. McNeal N, Scotti MA, Wardwell J, Chandler DL, Bates SL, Larocca M, et al. Disruption of social bonds induces behavioral and physiological dysregulation in male and female prairie voles. Auton Neurosci 2014 Feb;180:9-16. [doi: 10.1016/j.autneu.2013.10.001] [Medline: 24161576]
20. Thayer JF, Hansen AL, Saus-Rose E, Johnsen BH. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. Ann Behav Med 2009 Apr;37(2):141-153. [doi: 10.1007/s12160-009-9101-z] [Medline: 19424767]
21. Kaye WH, Wagner A, Fudge JL, Paulus M. Neurocircuity of eating disorders. Curr Top Behav Neurosci 2011;6:37-57. [doi: 10.1007/7854_2010_85] [Medline: 21243469]

Abbreviations

BIS-11: Barratt Impulsiveness Scale
BMI: body mass index
BN: bulimia nervosa
CBT: cognitive behavioral therapy
fMRI: functional magnetic resonance imaging
HR: heart rate
MRI: magnetic resonance imaging
MSIT: multi-source interference task
STAI: State-Trait Anxiety Inventory
VGT: video game therapy

Edited by G Eysenbach; submitted 15.01.14; peer-reviewed by U Schmidt, K Tchanturia, R Rodgers, V Cardi; comments to author 26.04.14; revised version received 29.05.14; accepted 22.07.14; published 12.08.14.

Please cite as:
Fagundo AB, Via E, Sánchez I, Jiménez-Murcia S, Forcano L, Soriano-Mas C, Giner-Bartolomé C, Santamaría JJ, Ben-Moussa M, Konstantas D, Lam T, Lucas M, Nielsen J, Lems P, Cardoner N, Menchón JM, de la Torre R, Fernandez-Aranda F
Physiological and Brain Activity After a Combined Cognitive Behavioral Treatment Plus Video Game Therapy for Emotional Regulation in Bulimia Nervosa: A Case Report
J Med Internet Res 2014;16(8):e183
URL: http://www.jmir.org/2014/8/e183/
doi:10.2196/jmir.3243
PMID:25116416

Original Paper

How Feedback Biases Give Ineffective Medical Treatments a Good Reputation

Mícheál de Barra[1], PhD; Kimmo Eriksson[1], PhD; Pontus Strimling[1], PhD

[1] Centre for the Study of Cultural Evolution, Stockholm, Sweden

Corresponding Author:
Mícheál de Barra, PhD
Centre for the Study of Cultural Evolution
Wallenberg Laboratory
Stockholm, WC1E 7HT
Sweden
Phone: 46 7531327690
Fax: 46 7531327690
Email: mdebarra@gmail.com

Abstract

Background: Medical treatments with no direct effect (like homeopathy) or that cause harm (like bloodletting) are common across cultures and throughout history. How do such treatments spread and persist? Most medical treatments result in a range of outcomes: some people improve while others deteriorate. If the people who improve are more inclined to tell others about their experiences than the people who deteriorate, ineffective or even harmful treatments can maintain a good reputation.

Objective: The intent of this study was to test the hypothesis that positive outcomes are overrepresented in online medical product reviews, to examine if this reputational distortion is large enough to bias people's decisions, and to explore the implications of this bias for the cultural evolution of medical treatments.

Methods: We compared outcomes of weight loss treatments and fertility treatments in clinical trials to outcomes reported in 1901 reviews on Amazon. Then, in a series of experiments, we evaluated people's choice of weight loss diet after reading different reviews. Finally, a mathematical model was used to examine if this bias could result in less effective treatments having a better reputation than more effective treatments.

Results: Data are consistent with the hypothesis that people with better outcomes are more inclined to write reviews. After 6 months on the diet, 93% (64/69) of online reviewers reported a weight loss of 10 kg or more while just 27% (19/71) of clinical trial participants experienced this level of weight change. A similar positive distortion was found in fertility treatment reviews. In a series of experiments, we show that people are more inclined to begin a diet with many positive reviews, than a diet with reviews that are representative of the diet's true effect. A mathematical model of medical cultural evolution shows that the size of the positive distortion critically depends on the shape of the outcome distribution.

Conclusions: Online reviews overestimate the benefits of medical treatments, probably because people with negative outcomes are less inclined to tell others about their experiences. This bias can enable ineffective medical treatments to maintain a good reputation.

(J Med Internet Res 2014;16(8):e193) doi:10.2196/jmir.3214

KEYWORDS

bias; social media; behavioral sciences; reputation systems; cultural evolution

Introduction

Across cultures and throughout human history, people have sought to alleviate suffering, shorten disease, and alter biological processes using medical treatments. An interesting feature of many medical treatments is that they are not directly beneficial; some even cause significant harm. This is true of Western folk beliefs, alternative medicines [1,2], traditional medicines [3,4], and historical "establishment" medicine like bloodletting [5]. It is also likely to be true of some contemporary medical treatments [6-8]. Treatments may be harmful either to the patient directly or cause harm because they replace other effective treatments, or result in broader environmental harms, as in the case of drugs derived from endangered species [2-4].

Medical treatments are very much cultural traits: rather than being invented anew by each individual, they spread from person to person through cultural processes. The prevalence of poor medical treatments is an anomalous outcome of cultural evolution because culturally acquired information in other domains of life is generally reliable and beneficial. Indeed, the extraordinary ecological success of the human species is, in part, due to our reliance on adaptive cultural information [9]. It is clearly true that humans routinely use cultural information to solve complex problems that, like medicine, entail delayed and/or stochastic feedback. The adaptive value of cultural information is thought to result from a number of mechanisms, such as learning heuristics whereby people selectively imitate more successful people, filtering whereby people evaluate the quality of socially acquired traits through experimentation, and natural selection whereby people with more beneficial cultural traits have more children who then learn these traits [10-12].

Some traditional medicines did have a direct benefit for the patient. Effective variolation, for example, was surprisingly common. For example, Yorba healers in West Africa carried smallpox scabs that could be used to induce a non-lethal infection and resultant immunity [13]. A number of vaccination techniques were being employed in 17th century India and China, and Edward Jenner's vaccination was long a part of English folk medicine [14]. Some globally important pharmaceutical products have their origins in traditional medicine; Artemisinin, a key anti-malaria drug, was part of ancient Chinese medicine [15]. Moreover, medicine—be it allopathic, traditional, or ancient—is not just about altering the course of disease. Medical experts will often have seen many people with similar diseases and thus they can help patients to understand what their illness is (diagnosis) and how it will play out over time (prognosis). For an anxious patient and his or her family, these are important services and they were probably carried out with some sophistication throughout history and across cultures. Moreover, by identifying and validating illness, medical experts may help the ill to garner social support and thus enable crucial rest and recuperation.

It is also clearly true that patients have undergone surgeries, ingested substances, and been subjected to a litany of other treatments with the explicit expectation that they would be helped. These expectations were not justified: the disease course was unaffected and/or the patient was directly harmed by the treatment. Ineffective treatments were common and remain common, and they warrant study [5]. Why then do harmful and non-beneficial medical treatments spread and persist?

We propose the following explanation. Irrespective of effectiveness, medical treatments typically result in a distribution of outcomes with some people improving, some deteriorating, and others experiencing little change. Suppose that the people who have more positive outcomes are more inclined to tell other people about their experience of the treatment than people who have poorer outcomes. This may occur because people recall their successes better than their failures, because people believe others' success stories, or because people are embarrassed to have adopted an ineffective treatment. Whatever the cause, such a bias would systematically distort the information available to other naive individuals who are seeking an effective treatment—the reputation of a treatment will exceed its real effect.

This hypothesis is assessed using a variety of methods. First, we compared clinical data on weight loss diets with weight loss reported in reviews of books on these diets. Reviews were taken from Amazon, a popular online marketplace where consumers can post reviews of products. We also made a similar comparison for unproven fertility treatments based on herbs and vitamins. In both cases, we predicted that people with positive outcomes are more inclined to post reviews. In a series of experimental studies, we then tested whether the bias of such reviews is sufficient to influence preferences for treatments. We predicted a preference for weight loss diets accompanied by typical reviews (as sampled from Amazon) over diets accompanied by undistorted reviews (ie, reviews that are representative of the diet's true effect obtained by purposefully sampling and/or editing of the review). Finally, we used a mathematical model to explore some implications of such reputational distortion.

Methods

Study 1: Weight-Loss Diets

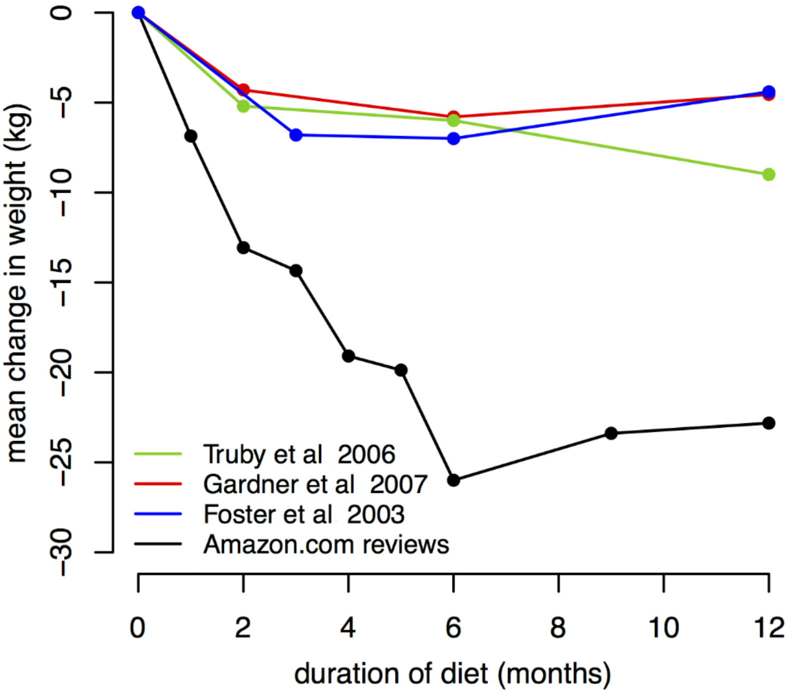
In order to make the Amazon and clinical data directly comparable, we made several assumptions and simplifications. Readers interested in conducting alternative analyses or comparisons can access the raw data and analysis syntax from the figshare data repository [16].

The Atkins Diet has been tested in several clinical trials and is the most commonly reviewed diet book on the Amazon online bookstore. We downloaded the 1359 reviews written on or before November 18, 2012. We extracted the duration of the diet and the total weight change from each diet review where this information was provided. If weight change at two time-points was mentioned (eg, 1 kg loss after 1 week and a 3 kg after 1 month), only the longer duration and associated weight change was recorded. If the review described the experiences of more than one person, only information about the author was recorded. If the review only discussed the weight change of a person besides the author, then that person's weight change was recorded. In total, 587 reviews included both a weight change and a time period over which this change occurred. The median diet duration was 42 days. To calculate an average weight loss at 1, 2, 3, 4, 5, 6, 9, and 12 months, we averaged the reports nearest each of these points in time. We excluded reviews of diets that lasted less than 2 weeks or more than 15 months.

The "true" effects of the Atkins diet were assessed using three clinical trials [17-19] in which participants received the Atkins diet book. In two of these trials [18,19], the intervention also entailed meeting a dietitian to discuss the diet and the participant's progress. Basic information about average weight loss in the Atkins diet arm could be extracted from the published manuscript, but to assess the distribution of outcomes, individual level data were needed. Only Gardner et al [18] were willing and able to share their raw data. The Gardner trial examined weight change among 311 premenopausal overweight and obese women, 77 of which were randomly allocated to the Atkins diet.

Participants received the Atkins book and met in groups of six once per week for 8 weeks to discuss the diet and book with a dietitian. Although Amazon reviewers are not all premenopausal women, Figure 1 shows that the average effect of the Atkins diet is broadly similar in several different populations. Moreover, given that the intervention involved reading the books *and* meeting with a dietitian, the clinical trial weight loss levels are likely to exceed that found in the general population. We compared the clinical weight change at 2, 6, and 12 months with Atkins reviews written between 1.5 and 2.5 months, 5 and 7 months, and 9 and 15 months respectively.

Figure 1. Average weight loss on Atkins diet reported in 3 clinical trials and Amazon reviews. Amazon data points were calculated by averaging reviews nearest the time points 1, 2, 3, 4, 5, 6, 9, 12 months. The numbers of reviews averaged to create the Amazon data points were 129, 60, 60, 23, 22, 19, 26, and 29 respectively.



Study 2: Fertility Treatments

On May 7, 2013, reviews of FertilAid (n=206), Fertilitea (n=198), and FertilityBlend (n=80) were downloaded from Amazon.com and reviews of Pregnancycare (n=68) were downloaded from Amazon.co.uk (total N=552). These are the most commonly reviewed herbal/vitamin pregnancy pills on Amazon.com and Amazon.co.uk. The following information was extracted from each review, if available: pregnancy status, length of time trying to conceive (TTC) while using the treatment and the length of time TTC before beginning the treatment, presence/absence of a previous pregnancy, the woman’s age, the man’s age, the presence/absence of polycystic ovary syndrome (PCOS), and presence/absence of past pregnancy. Reviews were excluded if the author explicitly stated that pregnancy was not the desired outcome of the treatment.

There is no strong evidence that these treatments enhance fertility in the general population. One pilot study found Pregnancycare was associated with higher pregnancy rates in subfertile/infertile women undergoing ovulatory induction [20] but none of the Pregnancycare reviewers on Amazon reported using Clomid or other ovulatory induction treatments. Another low-power study reports higher pregnancy rates among 53 FertiliBlend users who had previously tried to conceive for 6-36 months [21] but, in the absence of follow-up studies with greater power, it is difficult to ascertain if this difference between treatment groups was clinically meaningful. The National Institute for Health and Care Excellence (NICE) does not recommend any of the aforementioned treatments and notes that “the effectiveness of complementary therapies for fertility problems has not been properly evaluated” [22]. Given the paucity of rigorous data, we assume that these treatments have little effect on fertility.

The pregnancy rates reported on Amazon were compared to pregnancy rates in a prospective study of conception risk in 346 German women [23]. Specifically, pregnancy rates were extracted from data used to generate the Kaplan-Meier survival curves in Figure 1 of that study. The Kaplan-Meier curve corrects for biases due to participant dropout and is considered a best estimate of true pregnancy rate. If women are more likely to write a review after a positive outcome (that is, pregnancy), then conception rates reported in Amazon should be higher than conception rates in the prospective trial. Several important differences between the prospective study and the Amazon data should be noted. First, while the prospective study reports duration TTC in number of cycles, most reviewers report time TTC in days, weeks, or months. Menstrual cycle lengths are quite variable [24] but to enable a direct comparison we assumed one cycle is equivalent to 28 days. Second, women in the prospective study were shown how to use temperature/cervical-mucus monitoring to ensure intercourse occurred on the most fertile days of the cycle. Third, cycles in which intercourse did not occur during fertile days (3%) were excluded from the analysis. Fourth, in the prospective trial, data collection commenced on the month that women switched from oral contraception to “fertility-focused intercourse”. In contrast, of the 153 Amazon reviewers who reported a pre-treatment period trying to conceive, the median period trying to conceive was 1 year. Just 8% of 340 women in the prospective study had not conceived within 12 cycles of fertility-focused intercourse [23]. This indicates that subfertility and infertility is more prevalent among the Amazon reviewers than in prospective study participants. A total of 38 of 558 reviewers (6.9%) reported PCOS, while 83 (14.9%) reported other fertility-related problems (eg, irregular cycles); couples with fertility problems were excluded from the prospective study. Because the prospective study entailed fertility education, exclusion of couples with fertility problems, and the exclusion of cycles where fertile-period intercourse did not occur, the reported conception rate is likely to be higher than what is found in the general population. The comparison between this prospective study and the Amazon reviews is therefore a conservative test of our hypothesis. We are aware of one factor that may bias the results in the other direction: only pregnancies confirmed by a clinician were recorded in the prospective study while any reported pregnancy was included in the Amazon reviews. However, modern digital home pregnancy tests are generally considered reliable.

Study 3: How Distorted Reputation Influences Treatment Choices

In a series of online experiments, participants recruited from Mechanical Turk, Amazon’s online crowdsourcing marketplace, were presented with two diets and a series of reviews and were then asked to choose between the diets. All participants resided in America, 61% were male and the mean age was 33 years (SD 11). The diet books were *Dr. Atkins Diet New Revolution* and *The 17-Day Diet*. All reviews were extracted from Amazon. Two sets of books/reviews were shown on different pages and the order of presentation was randomized. In one condition, the Atkins reviews were “undistorted” by (1) drawing the reviews from a population of reviews with 200 words or less and an average of 3.5 stars (SD 0.99), corresponding to the average and standard deviation satisfaction rating given to diets in a longitudinal study [25], and (2) adjusting the reported weight change to match the average loss at that time point in clinical trials (calculated using Figure 1). *The 17-Day Diet* reviews were selected randomly from reviews that explicitly stated a weight loss and duration and consisted of 200 words or less (mean number of stars 4.4, SD 0.99). In the other condition, *The 17-Day Diet* reviews had the reputational distortion removed using the same procedure (mean 3.5, SD 1.0), and the Atkins reviews were selected randomly from a sample of reviews that stated duration and weight loss (mean 4.4, SD 1.01). Thus, each book was shown alongside three reviews that were either randomly selected Amazon reviews or purposively selected and edited so as to be consistent with clinical findings. After reading the reviews, participants were asked, “Imagine you decide to begin a diet. Which of these two diets would you begin?”

Ideally, each participant would see a different selection of reviews randomly drawn from the appropriate population. However, technical constraints of our experimental software made this impossible and so instead we ran three versions of each experiment using different reviews randomly selected from the same population. We then averaged the results for these three versions. This procedure was intended to reduce the probability that chance properties of any one set of selected reviews would exert too much influence on the final result. The results were broadly similar across all three versions of the experiment. The results for each condition and the characteristics of the selected reviews are available in Multimedia Appendix 1. Experiment 2 followed the exact same procedure except the diets only differed in positivity—both sets of reviews reported a similar average weight loss. In Experiment 3, the diet reviews were similar in positivity (3.4 stars) but reported different average weight loss. In every case, the dependant variable was diet chosen.

The Act concerning the ethical review of research involving humans (2003:460) regulates research with human subjects in Sweden. Studies need approval only if personal data is collected (ie, race or ethnic origin, political opinions, religious or philosophical beliefs, or membership of a trade union, and data on health or sex life) or if there is an attempt to physically or mentally influence the participant. These studies do not meet these criteria. Participants were clearly informed that by submitting their responses to the questionnaire they consented to the responses being used for research.

Results

Study 1: Weight-Loss Diets

In the first study, we compared clinical data on weight loss diets with weight loss reported in reviews of books on these diets. Clinical trials indicate that the Atkins diet results in an average weight change of about −7 kg over the first 6 months and a regain of about 2 kg over the subsequent 6 months [17-19]. In Amazon reviews, the average weight change is about −25 kg after 6 months and −20 kg after 12 months. As Figure 1 shows, the average beneficial effect reported in reviews of the Atkins diet exceeds the real effect at all time points.

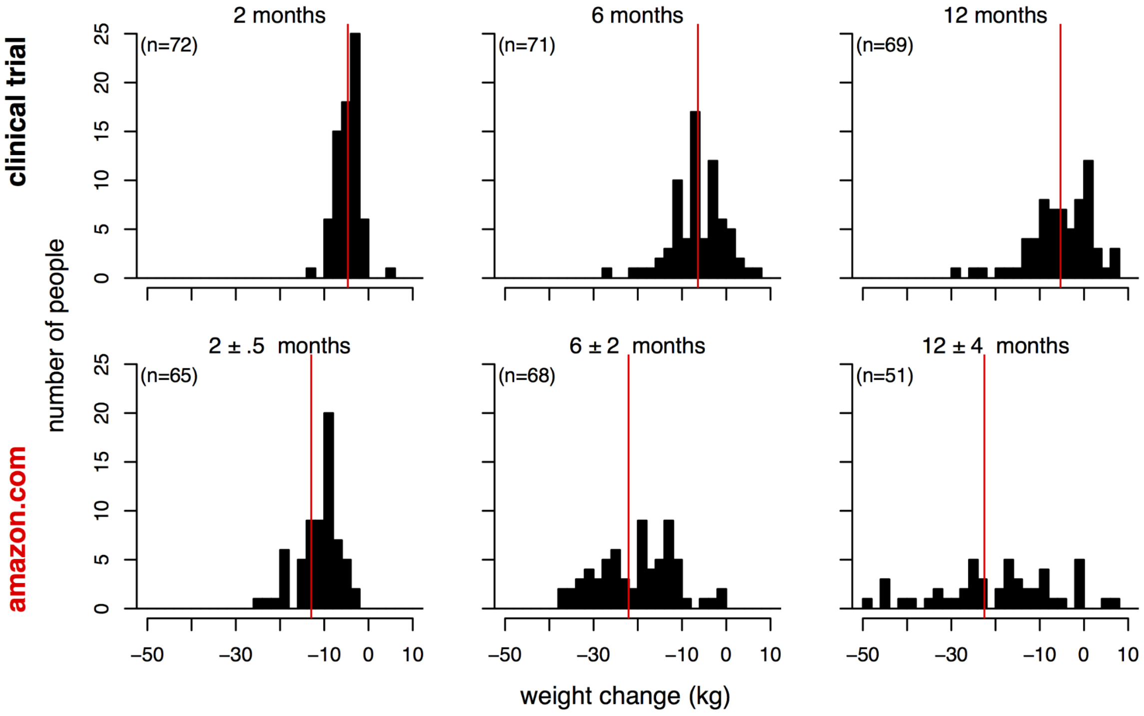
In Amazon reviews, weight loss is positively correlated with the number of stars (Spearman’s $\rho=.43$, $P<.001$), the diet duration ($\rho=.71$, $P<.001$), the word count ($\rho=.14$, $P<.001$), the number of capitalized letters ($\rho=.1$, $P=.01$), but not with the number of exclamation marks ($\rho=.05$, $P=.2$).

Individual level data from a 2007 clinical trial by Gardner et al [18] enabled a detailed comparison of real and reputed effects at three points in time (see Figure 2). The difference between the review data and clinical data was statistically significant at 2 months ($t_{69.8}=5.63$, $P<.001$, Cohen’s $d=0.98$), 6 months ($t_{92}=8.72$, $P<.001$, $d=1.48$), and 12 months ($t_{60}=5.86$, $P<.001$, $d=1.14$). In the clinical trial, participants sometimes lost and then regained weight. The average maximum weight loss for participants in the Gardner trial was 8.33 kg (SE 0.67); this maximum weight loss is also substantially lower than average Amazon weight loss of duration 2 months or greater. These data indicate that while 93% (64/69) of online reviewers reported a weight loss of 10 kg or more, just 27% (19/71) of trial participants experienced a similar weight loss level.

It is possible that the difference between real and reputed weight loss results from fake reviews written by individuals with a vested interest in Atkins sales. Fake reviews are unlikely to be produced continuously over time or at a rate proportional to the number of real reviews. Instead, they should be clustered at strategic times (immediately after an edition of the book is released) or in the period soon after the fake reviews are contracted. Therefore, we examined if the distortion applies over all time periods (suggesting a psychological bias) or if it exists only at certain time periods (suggesting fake reviews drive the distortion). The sample was split into deciles. Each decile contained 50+ individuals, and the deciles spanned from 1996 to 2012. Using the data from Gardner et al, we calculated the predicted weight loss for each participant. Gardner et al provide weight measurement at four time points; weight loss was assumed to be linear between these points. The difference between predicted and actual weight loss was calculated for each participant. A series of 10 one-sample *t* tests showed that in every time period there was a statistically significant distortion (maximum *P* value=.00005). Moreover, the difference between the predicted and real weight loss was of a similar magnitude in each decile (minimum mean difference 6.12, average mean difference 7.56, SD 1.41).

The subset of reviews that include weight change and diet duration information were somewhat more positive than total sample of reviews (mean of 4.43 stars vs 4.06 stars). An alternative explanation for the deviation between the Amazon reviews and the clinical trials results is that people with negative outcomes are less inclined to include specific information about the weight change and duration. In Multimedia Appendix 2, we show that a similar pattern of results is seen when a subset of reviews with a star distribution that matches that of the total sample is analyzed. This alternative hypothesis can therefore be rejected.

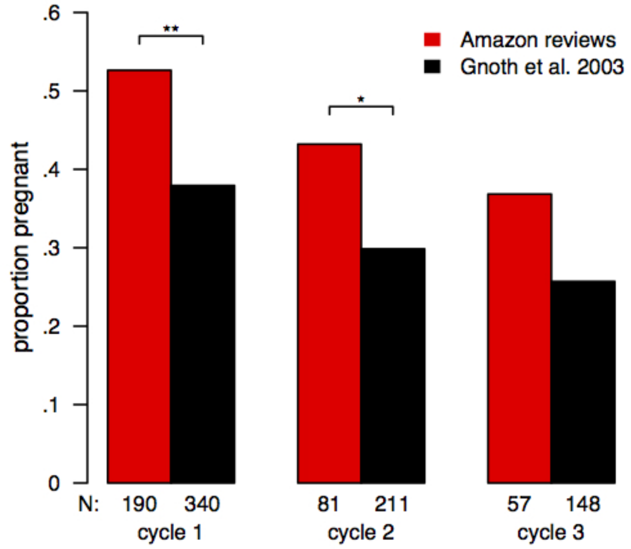
Figure 2. Comparison of weight loss distributions on Amazon reviews (bottom row) and clinical trial (11, top row) at three time points. Horizontal red lines indicate mean weight change. Outliers with weight loss >50 kg are not shown but are included in mean calculation.



Study 2: Fertility Treatments

In the second part of our study, we compared fertility data with Amazon reviews of unproven fertility treatments based on herbs and vitamins. In the 552 reviews analyzed, 186 people reported becoming pregnant after taking the treatment, 327 indicated they were not pregnant, and in 39 reviews it was unclear if a pregnancy occurred and/or the reviewer stated that pregnancy was not the desired outcome of the treatment. The duration of the medical treatment was stated in 443 reviews. Excluding the reviews where pregnancy was not reported/desired or the duration of the medical treatment was less than a week, 45.3% (173/382) reported becoming pregnant. Of the women who became pregnant, the median and mean time to pregnancy was 30 and 46 days, respectively. The mean time to pregnancy in the longitudinal study was considerably longer: 3.6 cycles or, if we assume a 28-day cycle, 101 days. Figure 3 illustrates the proportions of Amazon reviewers and study participants who became pregnant in each of the first three menstrual cycles. Chi-square tests indicate that more Amazon reviewers than study participants became pregnant in cycle 1 (100 of 190 vs 129 of 340, χ^2_1=10.04, *P*=.001) and in cycle 2 (35 of 81 vs 63 of 211, χ^2_1=4.70, *P*=.03). In cycle 3, the difference was not statistically significant (21 of 57 vs 38 of 148, χ^2_1=1.97, *P*=.16).

Figure 3. Proportion of non-pregnant women who conceived in each cycle of prospective study and in Amazon reviews of herbal/vitamin fertility treatments. Amazon proportions were calculated by collating reviews in which treatment was used for 28±14 days (cycle 1), 56±14 days (cycle 2), and 84±14 days (cycle 3). 1 star (*) and 2 stars (**) indicate statistically significant differences at P<.05 and P<.01 levels, respectively.

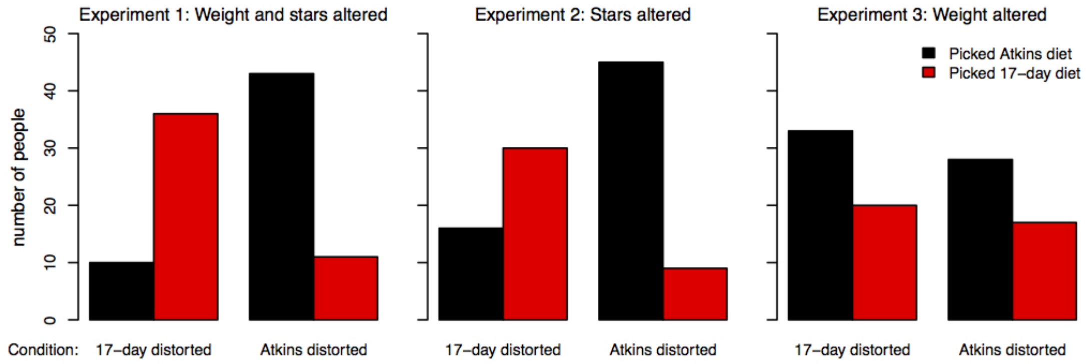


Study 3: How Distorted Reputation Influences Treatment Choices

Studies 1 and 2 demonstrate that the reputed benefits of medical treatments tend to exceed their actual benefits. The objective of Study 3 was to examine if this reputational distortion is large enough to influence people’s medical decision making.

Biased reporting can influence cultural evolution if the reputation of the treatment influences subsequent decision. We conducted three experiments with the objective of assessing how positively distorted sets of reviews might influence diet choice. Results indicated that participants were much more likely to pick a diet if its reviews were distorted with respect to both positivity (stars awarded to diet) and weight change (Experiment 1: $\chi^2_1=33.42$, n=100, *P*<.001) or distorted with respect to positivity alone (Experiment 2: $\chi^2_1=24.61$, n=100, *P*<.001). However, reviews that included distorted weight loss alone had no effect on preferences (Experiment 3: $\chi^2_1=0.02$, n=99, *P*=.89). These results are summarized in Figure 4.

Figure 4. Experiment 1 indicates participants prefer diet book with positive reviews and large weight loss over diet with positivity and weight change more representative of clinical trial results. Experiments 2 and 3 indicate that positivity alone but not weight change alone influence preferences.

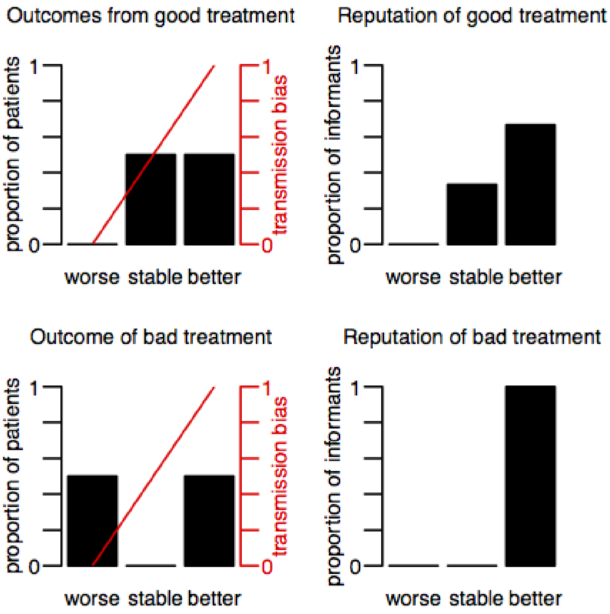


Mathematical Model

Can this mechanism account for the prevalence of harmful medical treatments across cultures? If the same kind of reporting bias affects all medical treatments, one might think that better treatments will still have a better reputation. However, this is not necessarily the case. Here, we show that the degree to which a treatment’s reputation is distorted by reporting bias will critically depend on the shape of the outcome distribution. In some circumstances, the result will be a superior reputation for an inferior treatment. The basic idea of the model is illustrated in Figure 5.

In order to isolate the effect of the reporting bias, we will make several strong assumptions about how well informed people are. First, we will assume that people have access to an infinite population of informants. These informants are honest, but they are more likely to share information if their outcome is better. Learners then choose the treatment with the best average reputation. This simple model shows that reporting bias can cause the spread of suboptimal treatments in a population.

Figure 5. Hypothetical example illustrating the effect explored in the mathematical model. With a reporting bias that makes bad outcomes unobservable, the poorer treatment obtains a better reputation (bottom row: all improve) than the good treatment (top row: 3/4 improve, 1/4 remain stable).



The specific assumptions of the model are as follows: for a focal treatment, let $d(x)$ denote the density function that describes the distribution of outcomes (measured on some scale of goodness). To implement a reporting bias such that a better outcome is always more likely to be reported than a worse outcome, we assume an individual who obtains outcome x will report this outcome with probability $f(x)$, where f is a strictly monotone increasing function of x. A learner has access to the reports of an infinite number of people who have tried the treatment in question. The learner then observes a distribution of reported outcomes with density function $d(x)f(x)$ divided by a constant factor $\int_{-\infty}^{\infty} d(y)\, f(y)\, dy$ to maintain unit total probability. Thus, the average observed outcome is shown in Figure 6.

Figure 6. Average observed outcome.

$$\frac{\int_{-\infty}^{\infty} d(x)f(x)x\ dx}{\int_{-\infty}^{\infty} d(y)\,f(y)\ dy}$$

To formalize comparison of treatments, define one treatment as *strictly better* than another treatment if the probability that it gives an outcome better than x is always at least as high, and for some x higher, than the probability that the other treatment gives an outcome better than x. It then holds that for any given treatment, one can always find another outcome distribution, corresponding to a hypothetical treatment, such that the former treatment is strictly better than the latter treatment but nonetheless the learner will choose the latter treatment because it will have a better average observed outcome.

We model goodness of outcomes as values on the real line. Reporting bias is modeled as a strictly monotonic function f satisfying $f(x) \rightarrow 0$ as $x \rightarrow -\infty$ and $f(x) \rightarrow 1$ as $x \rightarrow \infty$. Let $d_1(x)$ be the density function of a non-degenerate probability distribution on the real line, and let $D_1(x)$ denote its cumulative distribution function.

Theorem 1

For every distribution $d_1(x)$ with cumulative distribution function $D_1(x)$, there exists a distribution $d_2(x)$ with cumulative distribution function $D_2(x)$ that is strictly worse (ie, $D_2(x) \geq D_1(x)$ for all x and $D_2(x) > D_1(x)$ for some x) but is perceived as better using some perception bias function f. That is the average observed outcome of the strictly worse distribution $d_2(x)$ is better than the average observed outcome of $d_1(x)$ (Figure 7).

Figure 7. Equation shows treatment two appears more effective.

$$\frac{\int_{-\infty}^{\infty} d_2(y)f(y)\,y\ dy}{\int_{-\infty}^{\infty} d_2(y)f(y)dy} > \frac{\int_{-\infty}^{\infty} d_1(y)f(y)\,y\ dy}{\int_{-\infty}^{\infty} d_1(y)f(y)dy}$$

What the theorem says is that there exists a distribution $d_2(x)$ of outcomes that is strictly worse than $d_1(x)$, but that will nonetheless (under the reporting bias f) have higher perceived value (see Multimedia Appendix 3).

Discussion

Principal Findings

We found that the reputed benefit of weight loss diets and fertility treatments is larger than the real benefit, apparently because people with typical or poorer outcomes are less inclined to tell others about their experiences. Thus, the real-world reputation of medical treatments seems to be subject to a reporting bias akin to the publication bias toward positive results that is seen in scientific research [26]. Moreover, we found the resultant reputation distortion to be large enough to influence people’s decisions about which diet to begin.

An alternative explanation for the unduly positive reputation of the Atkins diet in our data is that reviewers make mistakes or lie. However, it seems unlikely that measurement error could account for the three- to four-fold difference in weight loss we observed, or that reviewers exaggerate to such a large degree in an online review. Similarly, error alone seems unlikely to account for the significant differences in conception rates, and reviewers had little motivation to lie about pregnancy status. It is also unlikely that fake reviews (written by people wishing to inflate or deflate the reputation of the product) account for our results. The deviation between the reputed benefits and the real effects of the treatments is similar across all eight Atkins diet durations (Figure 1), similar across 15 years of diet book reviews, and is similar over all three menstrual cycles. This consistent pattern of deviation seems more likely to stem from characteristics of human psychology than from deliberate fake review creation.

Although our analysis focused on specific weight change, the experimental data indicates that the general positivity of the review has a stronger influence than the reported weight loss. However, it is not crucial to our main hypothesis whether people are mainly influenced by the emotional or quantitative aspects of others' experiences because these are closely correlated, both in our data and in other studies of diet satisfaction and weight loss [27-29]. Our sample was perhaps less interested in losing weight than the population of people who are beginning diets. It is possible that prospective dieters would be more sensitive to specific weight information.

Conditions Where Reputation is Distorted

In summary, we found support for our hypothesis that ineffective and even harmful treatments may spread in a population when (1) treatments depend on word-of-mouth reputation, (2) treated individuals with poor outcomes can remain "invisible"' if they so wish, and (3) there is a broad range of outcomes. Moreover, the mathematical model shows that the distortion of reputations does not act equally across all treatments: a treatment that succeeds in pulling individuals from bad to intermediate outcomes may, paradoxically, seem worse than a treatment that fails to help individuals with bad outcomes. The bias may therefore account for the historical proliferation of ineffective medical treatments [5].

A slightly different, but conceptually similar, distortion may occur when doctors forget about patients who die under their care. Treatments like bloodletting are especially dangerous to individuals in poor health [30,31]. Given that such individuals were quite likely to remain sick or disabled for the remainder of their lives, a treatment like bloodletting may counter-intuitively appear effective because the past patients who have been bled appear healthier than the past patients who were never bled. What has really happened is that the doctor has "culled" the individuals most likely to remain ill or infirm. Patients killed by harmful treatments may be relatively easy to omit from considerations of treatment effectiveness simply because they have been removed from the community. Although the cause of distortion is different (patients with bad outcomes die and are forgotten versus patients with bad outcomes are inclined to remain silent), our mathematical model describes both cases.

It is not necessarily the case that treatments directly compete in the way our model assumes. Rather than comparing a number of treatments and selecting the one with the best reputation, people may simply adopt the first treatment that meets some criteria (eg, "two consecutive people rate it highly"). The reputational distortion we document means that such criteria will be met more frequently and thus it might cause people to adopt more treatments, including more ineffective ones.

More directly, this feedback bias may be one reason that people have unrealistically high expectations of weight loss diets and other medical treatments. For example, in a study where people were asked to estimate their "dream weight", "happy weight", "acceptable weight", and "disappointed weight", before they began a 48-week diet, 47% of participants did not even reach their "disappointed" weight [32]. Interestingly, participants' average "acceptable" weight change was very similar to the average weight change we found reported in Amazon reviews: a 25 kg loss.

This positive distortion in reputation has some important implications for the clinician. Patients are increasingly taking an active role in determining which treatments to adopt. It is unlikely that all the information used to make these decisions will come exclusively from medical professionals or rigorous research: people will listen to their friends, their family, and to other patients with similar experiences. Biases that undermine the reliability of this information, like the one documented here, will become increasingly important. Doctors and patients should be aware of them.

Conclusions

Researchers have pointed out that several processes make it very difficult to identify benefits and harms of medical treatments when data are not systematically collected. In particular, treatments with no direct effect will sometimes appear effective because of the statistical phenomenon known as regression to the mean and the physiological phenomenon known as the placebo effect [33,34]. It has also been suggested that treatments that prolong illness may, perversely, spread better because they are "demonstrated" for a longer period than effective treatments [35]. Here, we have explored an additional mechanism, reporting bias, and its logical consequence: when people with poor outcomes remain silent, the reputed benefit of a treatment will exceed its real effect.

Acknowledgments

We thank Daniel Cownden and Susanne Herbst for methodological suggestions, Alberto Acerbi for helping to extract the reviews from Amazon, Christopher Gardner for sharing diet data, and Erhard Godehardt and Christian Gnoth for the fertility data. The research was supported by the Swedish Research Council (grants 2009-2390 and 2009-2678). The funder played no role in this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Details of stimuli and results of experiments in Study 3.

[PDF File (Adobe PDF File), 173KB - jmir_v16i8e193_app1.pdf]

Multimedia Appendix 2

Analysis of a subset of reviews with star distribution equal to that of total review sample.

[PDF File (Adobe PDF File), 310KB - jmir_v16i8e193_app2.pdf]

Multimedia Appendix 3

Proof of mathematical theorem.

[PDF File (Adobe PDF File), 397KB - jmir_v16i8e193_app3.pdf]

References

1. Pittler MH, Brown EM, Ernst E. Static magnets for reducing pain: systematic review and meta-analysis of randomized trials. CMAJ 2007 Sep 25;177(7):736-742 [FREE Full text] [doi: 10.1503/cmaj.061344] [Medline: 17893349]
2. Ernst E, Singh S. Trick or treatment: the undeniable facts about alternative medicine. New York: W.W. Norton & Co; 2009.
3. Ayling J. What sustains wildlife crime? Rhino horn trading and the resilience of criminal networks. Journal of International Wildlife Law & Policy 2013 Jan;16(1):57-80. [doi: 10.1080/13880292.2013.764776]
4. Milliken T, Shaw J. The South Africa-Viet Nam Rhino Horn Trade Nexus. Traffic 2012:1-180.
5. Wootton D. Bad medicine: doctors doing harm since Hippocrates. New York: Oxford University Press; 2006.
6. Krumholz HM, Lee T. Redefining quality--implications of recent clinical trials. N Engl J Med 2008 Jun 12;358(24):2537-2539. [doi: 10.1056/NEJMp0803740] [Medline: 18539915]
7. Glasziou P, Moynihan R, Richards T, Godlee F. Too much medicine; too little care. BMJ 2013 Jul 02;347(jul02 2):f4247-f4247. [doi: 10.1136/bmj.f4247] [Medline: 23820022]
8. Wennberg JE. Tracking medicine: a researcher's quest to understand health care. New York: Oxford University Press; 2010.
9. Boyd R, Richerson PJ, Henrich J. The cultural niche: why social learning is essential for human adaptation. Proc Natl Acad Sci U S A 2011 Jun 28;108 Suppl 2:10918-10925 [FREE Full text] [doi: 10.1073/pnas.1100290108] [Medline: 21690340]
10. Enquist M, Ghirlanda S. Evolution of social learning does not explain the origin of human cumulative culture. J Theor Biol 2007 May 7;246(1):129-135. [doi: 10.1016/j.jtbi.2006.12.022] [Medline: 17275852]
11. Laland KN. Social learning strategies. Learn Behav 2004 Feb;32(1):4-14. [Medline: 15161136]
12. Richerson PJ, Boyd RW. Culture and the evolutionary process. Chicago: University of Chicago Press; 1988.
13. De Smet PAGM. Herbs, health & healers Africa as ethnopharmacological treasury. Netherlands: Afrika Museum; 1999.
14. Plotkin S, Plotkin S. A short history of vaccination. In: Plotkin SA, Orenstein WA, Offit PA, editors. Vaccines. Philadelphia: Elsevier Health Sciences; 2004.
15. Butler A, Khan S, Ferguson E. A brief history of malaria chemotherapy. J R Coll Physicians Edinb 2010 Jun;40(2):172-177. [Medline: 20695174]
16. de Barra M. Reputation of medical treatments: Dataset and analysis syntax. 2014. URL: http://figshare.com/articles/How_feedback_biases_give_ineffective_medical_treatments_a_good_reputation/843626 [accessed 2014-08-12] [WebCite Cache ID 6RmfLB5W9]
17. Truby H, Baic S, deLooy A, Fox KR, Livingstone MB, Logan CM, et al. Randomised controlled trial of four commercial weight loss programmes in the UK: initial findings from the BBC "diet trials". BMJ 2006 Jun 3;332(7553):1309-1314 [FREE Full text] [doi: 10.1136/bmj.38833.411204.80] [Medline: 16720619]
18. Gardner CD, Kiazand A, Alhassan S, Kim S, Stafford RS, Balise RR, et al. Comparison of the Atkins, Zone, Ornish, and LEARN diets for change in weight and related risk factors among overweight premenopausal women: the A TO Z Weight Loss Study: a randomized trial. JAMA 2007 Mar 7;297(9):969-977. [doi: 10.1001/jama.297.9.969] [Medline: 17341711]
19. Foster GD, Wyatt HR, Hill JO, McGuckin BG, Brill C, Mohammed BS, et al. A randomized trial of a low-carbohydrate diet for obesity. N Engl J Med 2003 May 22;348(21):2082-2090. [doi: 10.1056/NEJMoa022207] [Medline: 12761365]
20. Agrawal R, Burt E, Gallagher AM, Butler L, Venkatakrishnan R, Peitsidis P. Prospective randomized trial of multiple micronutrients in subfertile women undergoing ovulation induction: a pilot study. Reprod Biomed Online 2012 Jan;24(1):54-60. [doi: 10.1016/j.rbmo.2011.10.004] [Medline: 22138521]
21. Westphal LM, Polan ML, Trant AS. Double-blind, placebo-controlled study of Fertilityblend: a nutritional supplement for improving fertility in women. Clin Exp Obstet Gynecol 2006;33(4):205-208. [Medline: 17211965]

22. Fields E, Chard J, James D, Treasure T, Guideline Development Group. Fertility (update): summary of NICE guidance. BMJ 2013;346:f650. [Medline: 23427132]
23. Gnoth C. Time to pregnancy: results of the German prospective study and impact on the management of infertility. Human Reproduction 2003 Sep 01;18(9):1959-1966. [doi: 10.1093/humrep/deg366]
24. Treloar AE, Boynton R, Behn B, Brown BW. Variation of the human menstrual cycle through reproductive life. Int J Fertil 1967;12(1 Pt 2):77-126. [Medline: 5419031]
25. Baldwin AS, Rothman A, Jeffery R. Satisfaction with weight loss: examining the longitudinal covariation between people's weight-loss-related outcomes and experiences and their satisfaction. Ann Behav Med 2009 Dec;38(3):213-224 [FREE Full text] [doi: 10.1007/s12160-009-9148-x] [Medline: 20082163]
26. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet 1991 Apr 13;337(8746):867-872. [Medline: 1672966]
27. Clarke KK, Freeland-Graves J, Klohe-Lehman DM, Bohman TM. Predictors of weight loss in low-income mothers of young children. J Am Diet Assoc 2007 Jul;107(7):1146-1154. [doi: 10.1016/j.jada.2007.04.016] [Medline: 17604743]
28. De Vet E, Nelissen RM, Zeelenberg M, De Ridder DT. Ain't no mountain high enough? Setting high weight loss goals predict effort and short-term weight loss. J Health Psychol 2013 May;18(5):638-647. [doi: 10.1177/1359105312454038] [Medline: 22933574]
29. VanWormer JJ, Martinez AM, Cosentino D, Pronk NP. Satisfaction with a weight loss program: what matters? Am J Health Promot 2010 Mar;24(4):238-245. [doi: 10.4278/ajhp.080613-QUAN-92] [Medline: 20232605]
30. Wardle TD. Co-morbid factors in trauma patients. British Medical Bulletin 1999 Jan 01;55(4):744-756. [doi: 10.1258/0007142991902754]
31. Wutzler S, Maegele M, Marzi I, Spanholtz T, Wafaisade A, Lefering R, Trauma Registry of the German Society for Trauma Surgery. Association of preexisting medical conditions with in-hospital mortality in multiple-trauma patients. J Am Coll Surg 2009 Jul;209(1):75-81. [doi: 10.1016/j.jamcollsurg.2009.03.015] [Medline: 19651066]
32. Foster GD, Wadden TA, Vogt RA, Brewer G. What is a reasonable weight loss? Patients' expectations and evaluations of obesity treatment outcomes. J Consult Clin Psychol 1997 Feb;65(1):79-85. [Medline: 9103737]
33. Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: recent advances and current thought. Annu Rev Psychol 2008 Jan;59:565-590. [doi: 10.1146/annurev.psych.59.113006.095941] [Medline: 17550344]
34. Morton V, Torgerson DJ. Effect of regression to the mean on decision making in health care. BMJ 2003 May 17;326(7398):1083-1084 [FREE Full text] [doi: 10.1136/bmj.326.7398.1083] [Medline: 12750214]
35. Tanaka MM, Kendal JR, Laland KN. From traditional medicine to witchcraft: why medical treatments are not always efficacious. PLoS One 2009;4(4):e5192 [FREE Full text] [doi: 10.1371/journal.pone.0005192] [Medline: 19367333]

Abbreviations

PCOS: polycystic ovary syndrome
TTC: trying to conceive

Edited by G Eysenbach; submitted 30.12.13; peer-reviewed by J Kendal, A Kandler; comments to author 09.05.14; revised version received 23.05.14; accepted 14.07.14; published 21.08.14.

Please cite as:
de Barra M, Eriksson K, Strimling P
How Feedback Biases Give Ineffective Medical Treatments a Good Reputation
J Med Internet Res 2014;16(8):e193
URL: http://www.jmir.org/2014/8/e193/
doi:10.2196/jmir.3214
PMID:25147101

Viewpoint

Promoting Business and Entrepreneurial Awareness in Health Care Professionals: Lessons From Venture Capital Panels at Medicine 2.0 Conferences

Talya Miron-Shatz[1*], MA, PhD; Itamar Shatz[2*]; Stefan Becker[3,4], MD, MBA; Jigar Patel[5], PhD; Gunther Eysenbach[6,7], MD, MPH, FACMI

[1]Center for Medical Decidion Making, Ono Academic College, Kiryat Ono, Israel
[2]Tel Aviv University, Tel Aviv, Israel
[3]Department of Nephrology, University Duisburg-Essen, Essen, Germany
[4]Institute for Drug Safety, University Hospital Essen, Essen, Germany
[5]London Office, McKinsey Company, London, United Kingdom
[6]JMIR Publications, Toronto, ON, Canada
[7]University Health Network, Techna Institute, and University of Toronto, Toronto, ON, Canada
[*]these authors contributed equally

Corresponding Author:
Stefan Becker, MD, MBA
Institute for Drug Safety
University Hospital Essen
Hufelandstr 55
Essen, 45122
Germany
Phone: 49 201 723 83844
Fax: 49 201 723 5633
Email: Stefan.Becker@uk-essen.de

Abstract

There are few mechanisms that bring the academic and business worlds together in a way that would maximize the success of health technology (health tech) start-ups by increasing researchers' knowledge about how to operate in the business world. Existing solutions (eg, technology transfer offices and dual degree MD/MBA programs) are often unavailable to researchers from outside the institution or to those who have already completed their primary education, such as practicing physicians. This paper explores current solutions and offers a partial solution: include venture capital (VC) panels in medical conferences. These VC panels educate academics on 2 important and interconnected issues: how to "pitch" their ideas in the business world and what to consider when creating a company. In these sessions, academia-based start-up companies present their ideas before a VC panel composed of professional investors and receive feedback on their idea, business plan, and presentation techniques. Recent panel recommendations from Medicine 2.0 conferences fell into 7 categories: (1) the product, service, or idea you are developing into a company, (2) determine market forces and identify the target audience, (3) describe your competitive advantage, (4) the business plan, (5) current and future resources and capabilities, (6) legal aspects, and (7) general advice on the art of pitching. The academic and business literature validates many of these recommendations suggesting that VC panels may be a viable and cost-effective introduction to business and entrepreneurial education for physicians and other health care professionals. Panels benefit not only the presenting companies, but also the physicians, psychologists, and other health care professionals attending the session. Incorporating VC panels into academic conferences might also illuminate the need for incorporating relevant business training within academia.

(J Med Internet Res 2014;16(8):e184) doi:10.2196/jmir.3390

KEYWORDS

start-up; entrepreneurship; health technology; capital funding, telehealth, eHealth, mobile health, health technology, technology transfer; health 2.0; business pitch; entrepreneurship programs

Introduction

There is currently a disconnect between academia and business: researchers lack significant business training during their education. This gap in training, which stems from the focus on professional education (eg, in medicine), can impair academic researchers' potential integration into the world of entrepreneurship and business management because they lack the business training to know i how to build a healthy business model [1]. This lack of business training for health care professionals is neither new nor unknown. A study from 1993, for example, demonstrated that only 3% of young physicians (younger than 45 years) felt that they were well prepared to manage the business aspects of medical practice [2]. Today, psychologists, physicians, and other academics are increasingly developing interventions for health improvement and disease prevention, yet the leap into large-scale implementation of these interventions usually requires business knowledge. Without this knowledge, researchers are often unable to successfully develop their ideas commercially and they cannot manage to turn them into successful products or companies [3].

Our overwhelming sense that the field is alive with effective interventions that do not later translate to scalable products or services impelled the creation of the venture capital (VC) panel at the Medicine 2.0 conferences. The disconnect between academia and business (or "industry") can be demonstrated in 2 noticeable areas: the curricula in medical schools and the agendas of academic conferences. These 2 areas offer barriers, but also hidden opportunities; by modifying them to include the discussion of the integration between science and business, we can bridge the current gap and increase researchers' knowledge about how to operate in the business world. In this paper, we seek to cast light on the academia-business gap, illuminate existing solutions and limitations, and offer a partial remedy that provides business education in a nutshell. Additionally, we hope that this solution will make health care professionals realize what is missing in their training, and therefore will stimulate demand for changes within the medical curriculum and training process.

To highlight the knowledge gap, we first examined the curriculum of Harvard Medical School. Recently ranked as the number 1 medical school for research in the United States [4], Harvard Medical School does not currently require its graduates to take any business classes as part of their education [5]. Thus, even after completing the 4-year program, most certified physicians who graduate from Harvard Medical School have little to no formal business training. This may hamper their professional development when they begin actively leading research in the medical field, running private practices, or creating health-related start-ups (if they decide to do this) [6].

Second, we examined the agendas of several prominent academic conferences and found that they focused solely on science, ignoring business implementation altogether. Some notable examples are the International Federation of Fertility Societies and the American Society for Reproductive Medicine Joint Annual Meeting [7], the British Academy of Audiology Annual Conference [8], and the World Congress of the World Society for Pediatric Infectious Diseases [9]. None of these conferences offered any business-related sessions during their 2013 events. Interestingly, even one of the biggest international medical trade fairs, Medica in Dusseldorf, Germany, has not staged a VC panel session thus far [10]. Even conferences that are attuned to the issue of implementation, such as the NHS Health and Care Innovation Expo, showcase innovations helping visitors to bring about changes, improvements, and renewals within the NHS benefiting the whole community, but do not include a panel to inform medical entrepreneurs on how to bring their solutions to the stage where they can benefit the entire community [11]. A welcome exception is the Doctors 2.0 & You conference [12], which concentrates on understanding how physicians use new technologies, such as Web 2.0 and social media, and the impact of these latest technologies on the relation between physicians and patients, colleagues, industry, and the public sector. Perhaps because of its focus on integration with the industry, the first session is a start-up contest bringing together 7 companies from 5 countries working on diverse aspects of digital health. Attendees are expected from a range of industries, including the public, and physicians, professional and patient associations, pharmaceutical companies, governments, and insurance companies [12].

Interestingly, the American Medical Association (AMA) Accelerating Change in Medical Education Conference (held in Chicago on October 4 and 5, 2013) brought together almost 200 leaders in medical education from across the United States to discuss innovations needed to bridge the gap between the training of medical students and the needs of the health care system. Although there was an opportunity to learn about the grant projects supported by the Accelerating Change in Medical Education initiative, there was no VC panel [13]. Because conferences can instruct and set an agenda for a field, this is a missed opportunity.

The lack of business training has adverse effects on doctors and other health care professionals' forays into the world outside of medical school. A recent article featured on the Cancer Network website [14] showed that many doctors are unaware of the significance of having a proper business plan for their practice and are often unable to design one even if they do grasp its importance. Therefore, many struggle financially while running their practices. The detrimental effect of the lack of business education is exacerbated when it comes to more complex financial and business issues. Specifically, the chief executive officer (CEO) of a patient relationship management company [15] lists the most common reasons for the failure of health technology (health tech) start-ups: a lack of specific focus or adoption point, misunderstanding the consumers' willingness to pay for the service or how much effort they would be willing to expend to use it, requiring too much money for development of the product, having too complex an organizational structure, and lacking understanding of reimbursement dynamics. The same problems were raised in an article explaining why business modeling is crucial in the development of eHealth technologies [16] and in an article that discussed the importance of understanding business and economic strategies during the development of eHealth solutions [17].

Furthermore, a lack of understanding of business models reduces the ability of start-ups conceived in academia to receive funding for their development. The present paper focuses on fundraising from private sources (primarily from VC funds as discussed subsequently). However, the need to present a convincing case for the viability of an idea from a business perspective also applies when seeking to raise money from government sources such as federal grants [18], an important source of funding for health care start-ups [19]. Each source of funding has its own merits and is aimed at people and companies with different goals, although considerable overlap does exist between the 2 sources. Some examples are the emphasis on assembling a skilled team, showing the need for the proposed solution, and explaining why it should work [20-21]. In both cases, compelling arguments assist in securing funding. The primary difference is that federal grants are generally aimed at scientists who require additional funding to further their academic research in congruence with their university [22], whereas VC funding is aimed at companies looking to expand and explore commercial opportunities for profit [23]. Thus, the latter places more emphasis on larger growth or commercialization independent of a host academic institution.

All these issues share a single commonality: scientists lack a proper introduction into the intricacies of the business world and, therefore, risk being in a suboptimal position to develop their idea into a working marketable concept.

Existing Solutions for Bridging the Academic-Industry Gap

There are 4 main solutions currently in place that aim to minimize the adverse effects of the problem. These are technology transfer offices, entrepreneurship centers, specialized entrepreneurship programs, and medicine/business dual degree programs. However, none of these solutions will solve the problem entirely.

The first solution is the technology transfer offices (tech transfer) present in many universities, companies, and government organizations [24]. Their role is to identify which research has potential commercial interest and how to best develop and use it [25]. Although they serve an important purpose, many tech transfers do not comprehensively educate scientists about how the business world works [26-28]. Although they have a definite positive impact on research development, tech transfers are an incomplete solution because, in our opinion, many fail to give researchers the tools necessary for them to flourish and succeed in navigating the business aspects of the health care industry.

The second solution that sets out to deal with scientists' lack of business experience is the establishment of entrepreneurship centers in universities. These centers provide valuable support and training to aspiring entrepreneurs or researchers who are interested in learning more about the business world [29]. Unfortunately, although these centers provide obvious benefits, their greatest drawback lies in their locality because they are inherently limited in their ability to help anyone outside of the specific university in which they are set up. For example, the Global Consortium of Entrepreneurship Centers (GCEC), which is the premier organization to promote cooperation between entrepreneurship centers from different universities, is currently comprised of over 200 centers across the United States [30]. However, this is limited in scope because researchers from universities without these entrepreneurial centers rarely benefit from this sort of support.

The third available solution is specialized entrepreneurship programs that provide business education to scientists, such as the Stanford Summer Program on Bio-Entrepreneurship [31]. These entrepreneurship education and training (EET) programs teach scientists how to develop their research into a viable product or a functioning company. A quantitative review of all literature on the subject showed that EETs have a positive impact on entrepreneurial success [32]. The study found a statistically significant relationship between EET and entrepreneurship-related human capital assets (r=.217) and between EET and entrepreneurship outcomes (r=.159). More importantly, the study showed that the relationship between EET and entrepreneurship outcomes is stronger for academic-focused EET interventions (r=.238) than for training-focused EET interventions (r=.151), which emphasizes the importance of EET for academics. Again, the shortcoming is that EETs are a localized solution with limited coverage. Despite having a definite positive impact, EET programs cannot reach most health care professionals and researchers.

The fourth solution is a combined Doctor of Medicine (MD) and Master of Business Administration (MBA) program. These dual degree programs are designed with the goal of training physicians who are skilled in both medicine and business management. The integrated curriculum is designed in a way that strives to increase the drive, enthusiasm, and ambition of the degree candidates, containing the most important concepts from both fields: from strategy, finance, marketing, and economics on the business end to anatomy, physiology, biochemistry, and all other related core science disciplines of medicine [33,34]. Such programs are currently available in over 50 universities around the United States [35]. Dual degree programs are also effective in that students who participate in a dual degree program often perform better academically and have a higher degree of satisfaction with their studies than students who complete only an MBA or Doctor of Pharmacy (PharmD) program [36,37]. Although they offer the best and most extensive form of combined training (as far as receiving a business and a medical education goes), these dual degree programs suffer from a shortcoming similar to the one mentioned previously: anyone who did not study in such a program is unable to benefit from their existence. In addition, there is a scarcity of similar programs accessible to physicians during or immediately after residency training [38]. Lately, distance learning and online technology have permeated all levels of business education. However, most profiled programs so far are for general MBAs rather than combined MD/MBAs [39]. None of the courses featured by the Financial Times' Online MBA Listing 2014 focused on health care. Once again, the solution falls short because it reaches only a relatively small portion of the health care population.

For some start-ups, "incubators" may play an important role. These programs are designed to support researchers coming up with ideas by providing an array of business and services resources. Key to the success of such cooperations are powerful networks, in which all partners can trust. These are vital for bringing together know-how and venture capital. Yet, incubators are found outside of academic settings.

A Partial Solution and Potential Catalyst for Change: Venture Capital Panels in Medical Conferences

Having established the existence of a problem—the lack of business training for health care professionals—and the drawbacks of current solutions, we would like to propose an additional (although partial) solution, which overcomes the locality issue, namely VC panels hosted in medical and health care conferences. We are aware that VC panels cannot solve the problem entirely. In fact, anything short of making extensive business classes mandatory in medical school is unlikely to be a perfect solution. However, we believe that VC panels are a highly time- and cost-effective means of getting exposure to a broader sample of health care professionals. For many attendees, this can be their first substantial interaction with the business aspects of the research world. Deciding to attend a 90-minute session is not as big a time or financial commitment as deciding to enroll in an MBA degree, for instance. Because of this, VC panel sessions in academic conferences may attract people who are only in the early stage of considering business training or are exploring the relevance of the business world to their practice. Thus, these sessions can serve as a catalyst for creating demand for business education to be included in medical and other training and continued education programs.

Venture capital is funding provided to start-up companies. A VC fund receives equity in the company in return for its investment [40]; therefore, they tend to be long-term investments [41]. VC investments generally occur after a seed-funding round (used to start the business) has already taken place, although some funds also invest at the seed stage [42]. In 2010, there were 462 active (investing at least US $5 million) VC firms in the United States who invested approximately $22 billion into nearly 2749 companies, 1001 of these companies receiving funding for the first time [43]. Business factors, such as the potential for rapid return on investment and a credible business plan, are generally considered more important than product characteristics [44].

A VC panel is where companies and start-ups present their idea to venture capitalists in front of an audience and they are often included in industry conferences, events, and television shows, such as *Shark Tank* or its UK equivalent *Dragon's Den*. A number of prominent events developed in the United States over the past 7 years within the field of health information technology (IT). Examples are the Venture+ forum at the Health Information Management Systems Society (HIMSS) conference (Venture+ 2014: Health IT and Partnering Forums [45]), the Telemedicine Venture Summit at the conference of the American Telemedicine Association (American Telemedicine Association 2014 [46]), and the HealthTech Conference [47]. All combine educational components with possibilities for start-ups (between 10 and 45 companies) to present themselves to a panel. In 2013 and 2014, most topics revolved around mobile health (mHealth), in particular, patient-doctor communication. Benchmarking of the events is difficult because the number of applicants, growth attendance, and criteria for selection is not always made publicly available. One example comes from the HealthTech 2013 Conference, which hosted the "Grand Rounds Innovation Showdown." During this event, 10 start-up companies in the health industry (chosen out of more than 150 applicants) pitched their product or service to a group of judges, in front of a crowd of more than 400 health care executives, IT decision makers, venture capitalists, and members of the press [47]. Unfortunately, no reliable data exist in examining the extent to which VC panels have affected the development of companies in which they have invested. Similar events in Europe are relatively rare: The Charité Entrepreneurship Summit has only recently started focusing on IT (Charité Entrepreneurship Summit 2014 [48]). The biggest Medical IT conference, Connecting Healthcare IT (conhIT), has not offered VC panels thus far [49].

The benefits of including VC panels in academic conferences extend both to the companies presenting and to the audience. The companies receive invaluable feedback and get to practice "pitching," an essential skill in the business world [50] that is not a part of the academic training process. For the audience, the benefits include hearing about innovative new companies, learning from the feedback the companies receive, and becoming more familiar with pitches and company presentations. Panel members also benefit from an early glimpse at cutting-edge scientific developments and from exposure to existing and future academic entrepreneurs. Networking opportunities abound for all parties involved.

Venture Capital Panels at the 2012 and 2013 Medicine 2.0 Conferences

The Medicine 2.0 conference, established by Gunther Eysenbach in 2008, focuses on subjects such as digital disease detection, health information on the Web, and business models in a Web 2.0 environment [51]. This conference is perfectly positioned for beginning to bridge the gap between industry and academia, and for suggesting a new agenda. It showcases studies by researchers who either developed interventions for improving health and the transfer of health information, or are evaluating existing practices. In an era of burgeoning innovation and technological advancement in health care, there is great opportunity to marry the 2 sides. We propose to achieve this not only by introducing academics to investors, but also from providing academics with the knowledge and know-how of turning their validated ideas into businesses.

For the past 2 years (2012 and 2013), the Medicine 2.0 conference included a start-up panel organized and chaired by Professor Talya Miron-Shatz, a decision scientist, industry consultant, and CEO of CureMyWay, a behavior change start-up. During the panel sessions, companies conceived inside or alongside academic institutions presented their ideas to investors and other stakeholders, and received feedback that also served to inform the audience in attendance of the requirements of obtaining funding from such sources.

Members of the panels in 2012 and 2013 were seasoned investors: William Cowen of Long River Ventures, Boston; Joseph Kvedar of Health Partners, Boston; Jay Mohr of Locust Walk Partners, Boston; Jigar Patel of McKinsey & Company, London; Sid Thekkepat of m8capital, London; and Jack Young of Qualcomm Ventures, San Diego.

The companies that presented to the 2012 and 2013 VC panels had interesting and novel ideas in various stages of development. They ran the gamut from a person with an idea, 2 people developing a service, a company that had already established an impressive advisory board and raised funds, and many variations in-between. Their ideas included query engines for medical information, an online teenager community for maintaining a healthy body image, a system incorporating cell phone cameras with real-life Petri dishes to test water quality in Africa and elsewhere, a platform facilitating medical research, a health app, and a system for providing physicians with the most-read articles in their field. This suggests that Health 2.0 entrepreneurs can found companies based on a wide range of capabilities. The panel feedback reveals similarities in business needs, despite broad diversity in start-up topics.

For many researchers, the VC panel was an eye-opening first encounter with the business world. Therefore, we aggregated the feedback from the panels and compiled a list of the most critical pieces of information that the panelists related to companies. Entrepreneurs need to consider all the points mentioned subsequently when preparing a business presentation, but they are also crucial when developing the business idea and the company itself. In addition to this benefit, the feedback from the panels can help to outline and prioritize the subjects that entrepreneurial programs cover.

Table 1 lists the specific topics that companies were required to include in their pitch, with an example from a fictitious company. In this example, the fictitious company developed an apparatus for avoiding spillage when applying eye drops. Although the pitch was only 6-7 minutes long, presenting companies were required to cover all relevant topics.

The remainder of this paper outlines lessons learned from the VC panels, validates these lessons using current scientific and business literature, and discusses the potential implementation of VC panels as a partial yet scalable solution to health researchers' lack of familiarity with the business world.

Table 1. Topics to be addressed in a business presentation (pitch) for a hypothetical product to reduce eye drop spillage.

Topic	Example
The need or the problem	Patients applying eye drops spill 30% of the drops outside their eye.
The current state of affairs	Unless someone helps the patient, there is 30% spillage. No gadgets exist to solve the problem.
The company's solution	A mechanical device that is placed on the eye. The eyedrops bottle is placed in it. This ensures the bottle stays steady and there is less spillage.
Why the company's solution is better than other solutions	It is cheap to produce and therefore affordable, it minimizes spillage by 70%, it can be sterilized, and it requires no special skill to use.
The market	100 million people worldwide apply eyedrops at least once a day.
Monetization	The device will be distributed by medical insurers to ensure efficacy of eyedrops and reduce medication waste, which leads to repurchase or the device will be sold to directly to consumers.
Development phase: technologically	There is a currently a fully functional prototype.
Development phase: team	There is an ophthalmologist on board as a chief scientific officer, an engineer as a CEO, and 2 graduate engineering students on the development team.
Funding so far	An NIH grant of $300,000 for 1 year, borrowed $45,000 from friends and family, and received a $100,000 angel investment.
Business proposition for investors: how much the company is looking to raise and under what terms	Seeking $1,000,000 for a postinvestment evaluation of $3,000,000.

Venture Capital Panel Recommendations at the 2012 and 2013 Medicine 2.0 Conferences

The panelist comments (from 2012 and 2013) converged into 7 key areas, explained subsequently.

The Product, Service, or Idea You Are Developing Into a Company

Similar to the introduction section in a scientific paper, as the presenter you need to assume that the people you are presenting to are intelligent, but not necessarily familiar with the specific issue or field you are working on. Again, like an introduction section, presentations require that you cover certain points before describing your results—or product in the case of VC panels.

1. Background: What problem does your product solve? Describe the current state of affairs, such as the magnitude of the problem. For example, “100 million people apply eye drops each day. Studies show that 25% of the active material is lost due to improper application. This reduces the effectiveness of the drops, causing drug switches, unnecessary doctor visits, and a 12% increase in eye infections.” Note that the background is based on scientific findings, but is very succinct and presented using simple terms. Information to include encompasses several aspects of the product, which go beyond the technical description of how the product operates. Specifically, in order to convince investors of the potential success of the product, the company needs to make educated prediction regarding usage and acceptance of the product, by consumers (patients), as well as other stakeholders, such as insurers and physicians.
    - What is your solution (the product)? This should be a concise description that people from outside of the industry will be able to understand.
    - How will the costumer use the product? This ties in with the description of the product and shows what sort of a relationship the target costumer will have with the product.
    - Are people willing to pay for your product? This dovetails with the questions regarding the business plan, subsequently, and should be backed up with facts (eg, market research, surveys, and similar product histories).
2. Stakeholder analysis: What are the issues that matter to people who might later wish to use the solution and to those who would be willing to pay for it? Particularly in the complex health care arena, consideration needs to be given to any group or individual who can affect the achievement of your company’s objectives or is affected by them [52]. For example, “Health insurers are paying for spilled medication. They want to increase efficiency of application to reduce the need for repeated purchases of the drops. Insurers want patients’ health to improve or be steady because the insurer pays for additional treatments required due to deterioration. Patients suffer discomfort from spillage and from reduced effectiveness of the eye drops.”

Determine Market Forces and Identify the Target Audience

1. What market are you targeting and how big is it? This can be as specific as necessary to support the value of the product, but should be specific (eg, “Payers are spending US $20 million dollars in wasted eye drops each year”) rather than general (eg, “Health care in the United States is a US $3.8 trillion industry”).
2. Who are your competitors? In order to scan the competitive arena, you need to look beyond potential competitors and assess the competitive forces that can affect prospective profits [53]. A relevant question posed by panelists in this context is “What is your barrier to entry?” A *barrier to entry* is something that would stop your competition from developing a similar solution quickly and easily. A barrier could be an exclusive agreement you have already signed with major hospitals or health insurers, a patent, or anything else that requires ample time and/or money or other resources to develop, such as regulatory approval certification. Furthermore, you need to ask what degree your company is dependent upon suppliers and whether there are substitute offerings that could lure potential customers away.
3. Who is your customer? In the medical realm, customers can be divided into the 4 Ps: patient, provider, physician, and payer. There is also an important group that spans patient and provider that some products will directly target, namely caregivers. As a group, caregivers have a significant influence on the decisions individual patients make. In the previous example, the customer may be a pharmaceutical company that wants to differentiate its eye drops from others’ through using your device, a health insurer who wants to increase efficiency and reduce medication costs, or patients wishing to avoid the frustrating spillage. Likewise, an ophthalmologist can recommend the product to her patients to maximize efficiency and improve care.
4. What is your ability to ensure consumer engagement and loyalty? Engagement is increasingly becoming a parameter for evaluating companies that provide not just a service but also an experience to the user and should be quantified where possible.

Describe Your Competitive Advantage

Pitching without describing your competition, even briefly, is like writing an academic paper without citing any literature. Showing that competitors exist does not mean there is not room for your company. Rather, this is a positive because it indicates that a market exists for your product.

1. What are your competitors’ approaches to the problem? Are they currently successful (growth rate, revenue, etc)? What does their success/failure mean for you?
2. How are you different from other services? This is sometimes also referred to as your differentiator: the feature or element that will make customers choose you over the competition.
3. How sustainable is your competitive advantage? How quickly could competitors imitate your strategy? How quickly may resources become unavailable?

The Business Plan

The previous questions suggest that in order to pitch well, you need to be very familiar with the competition and to integrate these lessons into the building of your own product.

What is your revenue model (“show me the money”)? Including projected incomes and expenses, this is probably the biggest difference between the VC panel, which emphasizes financial sustainability, and the rest of the academic conference, which revolves around ideas, scientific findings, and implementation.

1. How are you going to make money or, in business jargon, to “monetize”? What is the payment model? If you plan to earn money primarily through reimbursement, does your model actually function? How long does it take to get paid?
2. Do you have an exit strategy? In other words, is there a feasible scenario for selling of your company or service

that would no longer require your involvement? This is something VC investors seek because they expect a high return on their investment.

3. What is the lifetime value of a customer versus the cost of recruiting a customer? The bigger the gap between the 2, the more lucrative your business proposition. This relates to the question of how long they will a customer/patient use the product.
4. How much capital have you raised so far and how? This includes any personal financial stake that you have in the company. For example, “We have already raised US $50,000 from personal savings and angel investment.”
5. How much capital do you need and how are you planning to raise it? You should be able to justify the required capital and be able to explain what you plan to spend it on over a given time period (eg, staff costs, patenting your ideas, developing a prototype, or expanding the business to other markets). For example, “We are looking for US $1.2 million to fund an 18 month rollout of our product to the top 20 payer systems in New England by recruiting a product manager, sales force, marketing department, and investing in research and development to improve product quality and reduce manufacturing costs.”
6. What business proposition are you looking to offer investors: how much money and under what terms? How much of your company are you willing to give up in exchange for the funding? It is crucial to be aware of this and have a plan before approaching the negotiations table.

Current and Future Resources and Capabilities

1. What phase is the company in technologically? How developed is the product? How far ahead are you in bureaucratic procedures such as patent filings and Food and Drug Administration (FDA) approvals?
2. What phase is the company in as far as a team is concerned? How experienced are team members? How well do they work together? Are they fully dedicated to the company (eg, what stake do they have in the company and what incentives do they receive)?
3. What are your monetary and development goals? What is the timeframe for the development of the business? Is it possible to accelerate progress using additional funding? Are there any potential bottlenecks that could hinder development? Are there any crucial deadlines?
4. Is your idea scalable and how? Scalability is the company’s ability to expand and deliver its products and services to multiple clients in various locations in a cost-efficient manner. In a digital world, this is simpler than it used to be. Scale, a prerequisite to growth, needs to be demonstrated.

Legal Aspects

1. How are you dealing with intellectual property laws? This is particularly pertinent to companies that evolved in a university setting, where the intellectual property often belongs to the institution, not the researcher.
2. How are you dealing with privacy laws? Data ownership needs to be established, as does adherence to regulations such as those determined by HIPPA [54]. As shown by Miron-Shatz and Elwyn [55], most patients will not be aware of breeches to the privacy of their data, but such breeches occur consistently. For example, if a company offers a platform where physicians can share pictures of various ailments (even if patient information is deindividuated, so they cannot be identified), its founders need to ensure that this is in compliance with the Health Insurance Portability and Accountability Act (HIPAA) and other regulations because patients have ownership of their own pictures, meaning these may not be able to be shared by others without clear permissions in place.

General Advice on the Art of Pitching

Apart from the content, the style and conveyance are also important, as business success hinges on impactful pitching. Similar to writing a scientific paper, having the data and the results is crucial, but the authors also need to present their arguments in a compelling manner so the journal accepts them (Textbox 1). Specific tips for presentation purposes were:

1. Be as focused and concise as possible. Both investors and your audience have a short attention span.
2. Use clear communication. Commercializing is a skill and has a language of its own. Beyond that, your communication needs to be clear and simple. Many comments revolved around the need to explain what the company does, from a number of angles, and in plain language. Make sure what you say is intelligible to people who are unfamiliar with the specific domain you operate in. On the first presentation slide, include a one-line description of your product/service (eg, “OpenTable for doctors”) so the panel and the audience will immediately know what your company does.
3. Use examples to highlight the need for your product and to show how you solve this need better, faster, and/or cheaper than anyone else does. You can do this by using cases of “the day in the life of...” a patient, physician, etc. This is the easiest way to show the panel how the product or idea works.
4. Show a lot of energy for your product to demonstrate that you believe in it and you will make it a success. This energy is something that a standard academic talk may lack because it is less of a “show” and more of a serious scientific presentation.

Textbox 1. Correlates of academic and business presentations.

One of the 2013 panelists, Jigar Patel, who has a PhD in computer science and artificial intelligence, related how he prepared for his first academic talk, some years ago. He mentioned preparing ferociously for hard questions he thought he might be asked, but spending too little time thinking about the story he was about to tell, recapping his results in a compelling manner.

This anecdote demonstrates that pitching is an acquired skill, which everyone, including those currently well versed in business lingo, had to learn and master at some point. This is similar to the challenge a company faces when pitching its business idea to prospective investors or business partners.

In a way, a business presentation is not different from writing the abstract of a scientific paper, which needs to convince its reviewers that it is worth publishing and its potential audience that it is worth reading. Just like a scientific paper, business presentations also have their logic and acceptable structure.

Final Words of Advice From the Panelists

1. Create barriers to entry by making it hard for others to imitate what you do. Accomplish this through a great user experience, intellectual property and patents, and/or through distribution channels and exclusive partnerships. The goal of this is primarily to protect yourself from intellectual theft of your product. This also makes it harder for other companies to compete with you directly by stealing your designs or methods.
2. Do not take it personally! Funders may choose not to invest just because they are in a late stage in the life of the fund, which means they are reserving money for continued investments in existing enterprises. There are many reasons why a funder may think your idea is brilliant, but still not invest.
3. Add value before seeking VC funding so you can retain more control in your company. Do so by looking for alternative sources of funding: collaborations, disease state groups, or nondilutive funding (eg, grants). These alternative sources of funding will likely also require a compelling business plan, pitch, or a proposal.
4. Practice makes perfect. Pitch to friends, colleagues, and mentors to get feedback before going to VCs. Consider filming yourself on video—this is a very honest way of realizing how you come across when you pitch. Get all the coaching and mentoring you can from people who will give you honest critical constructive feedback and give it your best shot.

Validating Venture Capital Panel Feedback Against Contemporary Business Advice

When examining the academic literature on the subject, we found that our advice to presenting companies about how to make a successful pitch and how to create a business plan was similar to that included in published business books and articles [56-61]. Business authors recommend defining the target market, identifying revenue mechanisms, and considering the competitive strategy. In addition, the strength of arguments is dependent on the passion, enthusiasm, credibility, interpersonal behavior, social signals, and honesty driving a fact-based presentation [60-64]. A number of articles from popular business magazines offering “golden rules” or “typical mistakes” resemble the advice we gave participants at the Medicine 2.0 VC Panel [65-69]. Among typical mistakes were not being concise during the pitch (eg, “the elevator pitch is longer than 1 minute” or “the PowerPoint presentation is too long”), not having a factually supported, well-written executive summary (which is a less-detailed version of a business plan), overlooking a realistic exit strategy for investors, and taking things personally (“failure to listen”) [67]. The “10 Tips Successful Business Pitch Presentation” on the Harvard Entrepreneurships website complements our panels’ conclusions [68]: “find the right investors to pitch to” and “let the investors ask themselves why they should join you.” Research findings that supplement our experience comes from a study that coded 11 episodes of the Dutch *Dragons’ Den* television show. During these episodes, 43 people pitched their new products to 5 investors. The author found that whether the language of the pitch was concrete or abstract did not impact investment decisions. However, pitchers who had more knowledge than what was included in the pitch, such as the market, target audience, and patents, had 6 times greater chance of receiving an investment [70]. It also verifies the need for a succinct presentation: “An investor pitch is a comprehensive plan that can be communicated according to the “rule of 3.” There are moments where you have to communicate your plan in less than 3 minutes.” [70]. This is especially relevant during the early stages of the pitching process. During the later stages of negotiations, times allow for 30 minutes or even 3 hours of presentation and discussion [70]. We can confidently conclude that the feedback from the panelists closely reflects advice from other existing business sources, meaning that VC panels are a credible means of educating academic entrepreneurs and would-be entrepreneurs in the workings of a pitch, the creation of a business plan, and the process of fund raising.

Conclusions

This paper has identified an inherent gap in business knowledge and training that may impede the translation of medical and psychological research into applied products, the commercialization of medical technologies, and the development of early stage health tech companies. We demonstrated that the gap reduces the chances of health care professionals engaging in medical start-ups and seeing their research insights implemented beyond the laboratory. Without business know-how, these professionals are less likely to successfully raise funds to support their companies and bring their ideas to fruition. The implications of this gap in knowledge go beyond the level of the individual academic entrepreneur and affect the entire health care industry. Medical and health care solutions developed in university settings can evolve into scalable intervention-based services and devices. Granted, there may be structural barriers to innovation and technology transfer, yet their focus would not necessarily be education. Hence, they were beyond the scope of this paper.

This paper discusses an additional solution: the inclusion of VC panels in medical conferences. These panels, which mirror a similar type of panel common in business conferences, have been organized at the Medicine 2.0 conference and have drawn considerable crowds and multiple submissions from companies, suggesting that all parties involved see potential gain in them and are willing to engage. The long-term effect of these panels can be evaluated by changes in the numbers of universities implementing existing solutions, in the generation of new solutions (mostly ones that overcome the locality issue), and in the ultimate creation of start-up companies in academia.

Similar to other solutions, VC panels are only a partial remedy to the lack of business knowledge of health care professionals. Rather than attempt to fix the problem in its entirety, VC panels can give both the companies presenting their products and the audience in attendance a chance to see how the business world functions. We regard the panels not only as a means of bridging the knowledge gap, but also as a way of sending a clear message to academicians and researchers: no matter how good your ideas are, you need to be able to understand how the business world works if you want to bring them to fruition. This can be an important teaching experience for health care professionals and researchers who are interested in developing products and services, and it is an experience that they are unlikely to receive anywhere else in the current medical educational system.

Acknowledgments

We would like to thank Gil Avnimelech, Ono Academic College, Kiryat Ono, Israel and Wolf-Thomas Nußbruch, Science Support Centre, University Duisburg-Essen, Germany, for valuable discussions. We acknowledge support by Deutsche Forschungsgemeinschaft for the publication costs.

Conflicts of Interest

One of the authors (GE) is editor of the Journal of Medical Internet Research (JMIR). Because of his involvement in the paper, assessment and peer review have been carried out entirely by an associate editor, who was not made aware of the fact that GE was a coauthor. GE has not been involved in any editorial decisions related to this paper. GE is also producer of the Medicine 2.0 Conference Series, and Chief Science Officer and co-founder of TrendMD, which was presented (by his co-founder) in one of the business panels.

References

1. Chen SH. The merging of medicine and management--McGill University's MD-MBA degree. Mcgill J Med 2006 Jan;9(1):3-4 [FREE Full text] [Medline: 19529800]
2. Cantor JC, Baker LC, Hughes RG. Preparedness for practice. Young physicians' views of their professional education. JAMA 1993 Sep 1;270(9):1035-1040. [Medline: 8350444]
3. Crutzen R. From eHealth technologies to interventions. J Med Internet Res 2012;14(3):e93 [FREE Full text] [doi: 10.2196/jmir.2050] [Medline: 22738742]
4. US News & World Report. 2013. Best medical schools URL: http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-medical-schools/research-rankings [accessed 2013-11-26] [WebCite Cache ID 6LQM1OOGL]
5. Harvard Medical School. 2013. Course Listings for Academic Year 2013- 014 URL: http://www.medcatalog.harvard.edu/ [accessed 2013-11-26] [WebCite Cache ID 6LQNA1Nx7]
6. Devji I. Cancer Network. 2012 Mar 13. The Business Plan: What Doctors Must Understand to Survive URL: http://www.cancernetwork.com/business-plan-what-doctors-must-understand-survive [accessed 2013-11-26] [WebCite Cache ID 6LQOkYMML]
7. American Society for Reproductive Medicine. 2013. 2013 Annual Meeting URL: http://www.asrm.org/IFFS-ASRM2013/ [accessed 2013-11-26] [WebCite Cache ID 6LQNnYhFW]
8. British Academy of Audiology. 2013. BAA 10th Annual Conference URL: https://www.eventsforce.net/FITWISE/media/uploaded/EVFITWISE/event_125/BAA%20Programme%202013%20FINAL.pdf [accessed 2013-11-26] [WebCite Cache ID 6LQNuXzfb]
9. World Society for Pediatric Infectious Diseases. 2013. 8th World Congress of the WSPID URL: http://www2.kenes.com/iser/scientific/Documents/WSPID%202013%20-%20Preliminary%20Scientific%20Program%20-%20NoSN.pdf [accessed 2013-11-26] [WebCite Cache ID 6LQO5o8iA]
10. MEDICA. Conferences at MEDICA URL: http://www.medica-tradefair.com/cipp/md_medica/custom/pub/content,oid,41239/lang,2/ticket,g_u_e_s_t/~/Conferences_Forums.htm [accessed 2014-03-07] [WebCite Cache ID 6NtiITLxt]
11. Health and Care Innovation Expo 2014. 2014. Annual Conference 2014 URL: http://www.healthcareinnovationexpo.com/index.asp [accessed 2014-05-15] [WebCite Cache ID 6PaUGQEuJ]
12. Doctors 2.0 & You. 2014. URL: http://www.doctors20.com/about-doctors20-and-you/ [accessed 2014-05-15] [WebCite Cache ID 6PaUVas2R]

13. American Medical Association. 2013 National Advocacy Conference URL: http://www.ama-assn.org/ama/pub/advocacy/events/national-advocacy-conference/2013-coverage-nac.page [accessed 2014-05-15] [WebCite Cache ID 6PaUTJQYd]
14. Devji I. Cancer Network. The Business Plan: What Doctors Must Understand to Survive URL: http://www.cancernetwork.com/business-plan-what-doctors-must-understand-survive [accessed 2013-11-26] [WebCite Cache ID 6LQOkYMML]
15. Chase D. TechCrunch. 2011 Aug 28. HealthTech FAIL: Lessons For Entrepreneurs From Health Startups Gone Awry URL: http://techcrunch.com/2011/08/28/healthtech-fail-lessons-for-entrepreneurs-from-health-startups-gone-awry/ [accessed 2013-11-26] [WebCite Cache ID 6LQOpHw4o]
16. van Limburg M, van Gemert-Pijnen JE, Nijland N, Ossebaard HC, Hendrix RMG, Seydel ER. Why business modeling is crucial in the development of eHealth technologies. J Med Internet Res 2011;13(4):e124 [FREE Full text] [doi: 10.2196/jmir.1674] [Medline: 22204896]
17. Scott RE, Mars M. Principles and framework for eHealth strategy development. J Med Internet Res 2013;15(7):e155 [FREE Full text] [doi: 10.2196/jmir.2250] [Medline: 23900066]
18. Avnimelech G, Teubal M. Targeting venture capital: lessons from Israel Yozma program. In: Bartzokas A, editor. Financial Systems, Corporate Investment in Innovation, and Venture Capital. Cheltenham, UK: Edward Elgar Publishing; 2004.
19. Frankel M, Chinitz D, Salzberg C, Reichman K. Sustainable health information exchanges: the role of institutional factors. Isr J Health Policy Res 2013;2(1):21 [FREE Full text] [doi: 10.1186/2045-4015-2-21] [Medline: 23692711]
20. Bauer DG. The ’How To’ Grants Manual: Successful Grantseeking Techniques for Obtaining Public and Private Grants. Lanham, MD: Rowman & Littlefield Publishers; 2011.
21. Hsu DH. Experienced entrepreneurial founders, organizational capital, and venture capital funding. Research Policy 2007 Jun;36(5):722-741. [doi: 10.1016/j.respol.2007.02.022]
22. Bozeman B, Gaughan M. Impacts of grants and contracts on academic researchers’ interactions with industry. Research Policy 2007 Jun;36(5):694-707. [doi: 10.1016/j.respol.2007.01.007]
23. Engel D, Keilbach M. Firm-level implications of early stage venture capital investment — An empirical investigation. Journal of Empirical Finance 2007 Mar;14(2):150-167. [doi: 10.1016/j.jempfin.2006.03.004]
24. Muscio A. What drives the university use of technology transfer offices? Evidence from Italy. J Technol Transf 2009 May 8;35(2):181-202. [doi: 10.1007/s10961-009-9121-7]
25. Caldera A, Debande O. Performance of Spanish universities in technology transfer: An empirical analysis. Research Policy 2010 Nov;39(9):1160-1173. [doi: 10.1016/j.respol.2010.05.016]
26. Association of University Technology Managers. About Technology Transfer URL: http://www.autm.net/Tech_Transfer.htm [accessed 2013-11-26] [WebCite Cache ID 6LQOtHNDL]
27. John Hopkins University. 2012. Technology Transfer: About JHTT URL: http://techtransfer.jhu.edu/about_us/ [accessed 2013-11-26] [WebCite Cache ID 6LQPEYzGJ]
28. Regents of the University of California. 2013. Technology Transfer Office: TTO Mission URL: http://invent.ucsd.edu/about/mission.shtml [accessed 2013-11-26] [WebCite Cache ID 6LQPIRxC8]
29. University of California San Francisco Office of Research. 2013. About Us URL: http://ita.ucsf.edu/about [accessed 2013-11-26] [WebCite Cache ID 6LQPQ1Vuq]
30. Global Consortium of Entrepreneurship Centers. 2006. URL: http://www.globalentrepreneurshipconsortium.org/index.cfm [accessed 2013-11-26] [WebCite Cache ID 6LQQt5Wrf]
31. Stanford School of Medicine. 2009. Bio-Entrepreneurship URL: http://med.stanford.edu/careercenter/resources/bioentrepreneurs.html [accessed 2013-11-26] [WebCite Cache ID 6LQPW911P]
32. Martin BC, McNally JJ, Kay MJ. Examining the formation of human capital in entrepreneurship: A meta-analysis of entrepreneurship education outcomes. Journal of Business Venturing 2013 Mar;28(2):211-224. [doi: 10.1016/j.jbusvent.2012.03.002]
33. Harvard Business School. 2013. MD/MBA Program URL: http://www.hbs.edu/mba/academic-experience/joint-degree-programs/Pages/harvard-medical-school.aspx [accessed 2013-11-26] [WebCite Cache ID 6LQQ05hXG]
34. Tufts University. 2013. MD/MBA Four Year Program URL: http://md.tufts.edu/Education/Combined-Degree-Programs/MDMBA [accessed 2013-11-26] [WebCite Cache ID 6LQQ46q75]
35. Association of MD MBA Programs. 2013. MD MBA Programs URL: http://mdmbaprograms.com/5.html [accessed 2013-11-26] [WebCite Cache ID 6LQQBPDyk]
36. Chumney EC, Ragucci KR. Student satisfaction and academic performance in a dual PharmD/MBA degree program. Am J Pharm Educ 2006;70(2):70-80.
37. Chumney EC, Ragucci KR, Jones KJ. Impact of a dual PharmD/MBA degree on graduates academic performance, career opportunities, and earning potential. Am J Pharm Educ 2008;72(2):26.
38. Baker S, Daginawala N. Leadership training for radiologists: a survey of opportunities and participants in MBA and MPH programs by medical students, residents, and current chairpersons. J Am Coll Radiol 2011 Aug;8(8):563-567. [doi: 10.1016/j.jacr.2011.02.013] [Medline: 21807350]
39. Financial Times. Online MBA Listing 2013 URL: http://rankings.ft.com/businessschoolrankings/online-mba-listing-2013 [accessed 2014-03-06] [WebCite Cache ID 6NsLEDolc]

40. Berger AN, Schaeck K. Small and medium-sized enterprises, bank relationship strength, and the use of venture capital. Journal of Money, Credit and Banking 2011;43(2-3):461-490. [doi: 10.1111/j.1538-4616.2010.00381.x]
41. Puri M, Zarutskie R. On the life cycle dynamics of venture ‐ capital ‐ and non ‐ venture ‐ capital ‐ financed firms. The Journal of Finance 2012;67(6):2247-2293. [doi: 10.1111/j.1540-6261.2012.01786.x]
42. Investopedia. Seed Capital Definition URL: http://www.investopedia.com/terms/s/seedcapital.asp [accessed 2013-11-26] [WebCite Cache ID 6LQQFRhcc]
43. National Venture Capital Association. Frequently Asked Questions URL: http://www.nvca.org/index.php?Itemid=147 [accessed 2013-11-26] [WebCite Cache ID 6LQQKhdB1]
44. Rea RH. Factors affecting success and failure of seed capital/start-up negotiations. Journal of Business Venturing 1989 Mar;4(2):149-158. [doi: 10.1016/0883-9026(89)90028-1]
45. Venture+. Health IT and partnering forums URL: http://www.ventureplusforum.com/ [accessed 2014-05-01] [WebCite Cache ID 6PF5UWphA]
46. American Telemedicine Association. Telemedicine Venture Summit URL: http://www.americantelemed.org/ata-2014/telemedicine-venture-summit [accessed 2014-05-01] [WebCite Cache ID 6PF5AQUAv]
47. VentureBeat, 2013. Innovation Showdown URL: http://venturebeat.com/events/healthbeat2013/innovation-competition/ [accessed 2013-11-26] [WebCite Cache ID 6LQPvfp3k]
48. Stiftung Charité. 7th Charité Entrepreneurship Summit URL: http://www.charite-summit.de/ [accessed 2014-05-01] [WebCite Cache ID 6PFJKkiRy]
49. conhIT. Connecting Healthcare IT URL: http://www.conhit.de/en/ [accessed 2014-05-01] [WebCite Cache ID 6PFINOFK0]
50. Williams A. A Study on the Art and Science of Pitching New Businesses [master's thesis]. Boston, MA: Massachusetts Institute of Technology; 2013.
51. Medicine 2.0. Overview URL: http://www.medicine20congress.com/ocs/index.php/med/med2014/schedConf/overview [accessed 2013-11-26] [WebCite Cache ID 6LQPjidO5]
52. Freeman RE, Harrison JS, Wicks AC, Parmar BL, De Colle S. Stakeholder Theory: The State of the Art. Cambridge: Cambridge University Press; 2010.
53. Porter ME. Harvard Business Review. 2008. The Five Competitive Forces That Shape Strategy URL: http://hbr.org/2008/01/the-five-competitive-forces-that-shape-strategy/ar/1 [WebCite Cache ID 6NyWO5Pak]
54. US Department of Health & Human Services. Summary of the HIPAA Privacy Rule URL: http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/ [accessed 2013-11-26] [WebCite Cache ID 6LQQg2uvL]
55. Miron-Scahtz T, Elwyn G. To serve and protect? Electronic health records pose challenges for privacy, autonomy and person-centered medicine. IJPCM 2011 Jun 10;1(2):405. [doi: 10.5750/ijpcm.v1i2.84]
56. Chesbrough H, Rosenbloom RS. The role of the business model in capturing value from innovation: evidence from Xerox Corporation’s technology spin ‐ off companies. Industrial and Corporate Change 2002;11(3):529-555. [doi: 10.1093/icc/11.3.529]
57. Abrams R. Successful Business Plan: Secrets & Strategies. Palo Alto, CA: The Planning Shop; 2003.
58. Zimmerer T, Scarborough NM, Wilson D. Essentials of Entrepreneurship and Small Business Management. Upper Saddle River, NJ: Prentice Hall; 2002.
59. Mullins JW, Komisar R. Getting to Plan B. Boston, MA: Harvard Business Press; 2009.
60. Chesbrough H. Business model innovation: it’s not just about technology anymore. Strategy & Leadership 2007;35(6):12-17.
61. Jordan JF. Innovation, commercialization, and start-ups. In: Schulz R, editor. Quality of Life Technology Handbook. Boca Raton, FL: CRC Press; 2014:279-299.
62. Jourdan Jr LF. The relationship of investor decisions and entrepreneurs’dispositional and interpersonal factors. Entrepreneurial Executive 2012;17:49.
63. Boni, Arthur A. The pitch and business plan for investors and partners. Journal of Commercial Biotechnology 2012;18(2):38.
64. Loeffler M. The Holistic Social Business Movement in Caldas (HSBM), Colombia. In: Grove A, Berg GA, editors. Social Businesses: Theory, Practice, and Clinical Perspectives. New York: Springer; 2014:45-61.
65. Frasch R. Forbes. 2013 Jul 09. 8 Mistakes Entrepreneurs Make When Pitching To Investors URL: http://www.forbes.com/sites/allbusiness/2013/07/09/8-mistakes-entrepreneurs-make-when-pitching-to-investors/ [accessed 2013-11-26] [WebCite Cache ID 6LQQNcg3L]
66. HarvardEntrepreneurships. 10 Tips Successful Business Pitch Presentation URL: http://www.harvardentrepreneurship.org/325/10-tips-successful-business-pitch-presentation.html [accessed 2013-11-26] [WebCite Cache ID 6LQQUhxmy]
67. Chen XP, Yao X, Kotha S. Entrepreneur Passion And Preparedness In Business Plan Presentations: A Persuasion Analysis Of Venture Capitalists' Funding Decisions. Academy of Management Journal 2009 Feb 01;52(1):199-214. [doi: 10.5465/AMJ.2009.36462018]
68. Henderson KJ. Chron. 15 Minute Business Presentation Tips URL: http://smallbusiness.chron.com/15-minute-business-presentation-tips-8.html [accessed 2013-11-30] [WebCite Cache ID 6LWVHB3N8]
69. Getty PM. The 12 Magic Slides: Insider Secrets for Raising Growth Capital. New York: Apress; 2013.
70. Beynen M. Dragons’ Den: How to Convince an Investor? The Effect of Concrete and Abstract Language on Success of a Pitch [master's thesis]. Netherlands: University of Leiden; 2013.

Abbreviations

AMA: American Medical Association
CEO: chief executive officer
EET: entrepreneurship education and training
FDA: Food and Drug Administration
GCEC: Global Consortium of Entrepreneurship Centers
HIMSS: Health Information Management Systems Society
HIPAA: Health Insurance Portability and Accountability Act
MBA: Master of Business Administration
VC: venture capital

Edited by P Wicks; submitted 12.03.14; peer-reviewed by H Potts, E Stanton, R Dorsey; comments to author 19.04.14; revised version received 07.05.14; accepted 27.05.14; published 06.08.14.

Please cite as:
Miron-Shatz T, Shatz I, Becker S, Patel J, Eysenbach G
Promoting Business and Entrepreneurial Awareness in Health Care Professionals: Lessons From Venture Capital Panels at Medicine 2.0 Conferences
J Med Internet Res 2014;16(8):e184
URL: http://www.jmir.org/2014/8/e184/
doi:10.2196/jmir.3390
PMID:25100579

Letter to the Editor

The Behavioral Intervention Technology Model and Intervention Mapping: The Best of Both Worlds

Rik Crutzen[1], PhD

CAPHRI, Department of Health Promotion, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Rik Crutzen, PhD
CAPHRI
Department of Health Promotion
Maastricht University
PO Box 616
Maastricht, 6200 MD
Netherlands
Phone: 31 433882828
Fax: 31 433671032
Email: rik.crutzen@maastrichtuniversity.nl

Related Article:

Comment on: http://www.jmir.org/2014/6/e146/

(J Med Internet Res 2014;16(8):e188) doi:10.2196/jmir.3620

KEYWORDS

mHealth; eHealth; behavioral intervention technology; intervention mapping

Mohr and colleagues recently presented an important paper describing an integrated conceptual and technological framework for eHealth and mHealth interventions, using existing models as their point of departure [1]. They mentioned that the focus of existing *psychological* models on clinical outcomes instead of behavior is one of two limitations. However, behavior is a central element in both of the provided examples—the theory of planned behavior and the social cognitive theory. The second limitation was that “psychological models do not include critically important factors that can guide the design and specifications for a BIT (behavioral intervention technology)” [1]. The idea of describing a BIT model that supports the translation of intervention components into features is accepted with open arms.

The BIT model was developed based on a review of three *design* models proposed by Ritterband [2], Fogg [3], and Oinas-Kukkonen [4], and addresses the limitations of these existing design models. Although the authors of the paper acknowledged that the review was not exhaustive, consideration of the intervention mapping (IM) protocol [5,6] would have added value to this viewpoint paper. The limitations of the existing design models are also addressed within IM, and there is overlap between the BIT model and IM, but each has their own unique contributions as well [7].

IM is a protocol for developing theory- and evidence-based interventions. IM describes the development process in six steps: (1) needs assessment, (2) specifying performance objectives and change objectives, (3) selecting theory-based intervention methods and practical applications, (4) designing and organizing the intervention, (5) specifying adoption and implementation plans, and (6) generating an evaluation plan. The protocol guides developers through each of these steps by means of specific tasks. I will focus on steps (2)-(4), as these steps are related to the limitations highlighted in the viewpoint paper by Mohr et al.

In step 2 of IM, the required actions (ie, performance objectives) for the behavioral outcomes have to be specified. Performance objectives are specific sub-behaviors of the health promoting behavior. For example, when you like to promote condom use (health promoting behavior), you want people to obtain condoms in order to use them. This is in line with the BIT component *Aims*. In the BIT model, attention is being paid to usage aims. Although use can be considered as adoption and implementation at the individual level [8], it is good to think ahead and consider use as a behavior with its own determinants [9]. Therefore, in IM, each performance objective is crossed with its determinants, resulting in the formulation of change objectives. These are specific goals of an intervention—to change the determinants of sub-behaviors.

In step 3 of IM, methods and applications are selected to influence the determinants in the desired direction. This step circumvents the limitations of existing design models in that “the Ritterband model does not articulate how technological components might be mapped onto more specific (and promixal) intervention goals” and the Fogg Behavior Model “does not purport to guide applications focused on changing attitudes or cognitions” [1].

A theoretical method is a general technique or process for influencing changes in the determinants, whereas an application is a specific technique for the practical use of theoretical methods in the context of the intervention. This is reflected in the BIT components *Behavior change strategies* and *Elements*, but parameters for use are not mentioned in the BIT model. The parameters for use are conditions that need to be met for a practical application to accurately reflect the theoretical method. If these parameters are lost in translation from method to application, then the method may not be used correctly and its effectiveness might be undermined. For example, feedback as a method works well if the feedback is personalized, follows behavior in time, and is specific [10]. This is different from the method of providing information about others’ approval, which can only work optimally if positive expectations are available in the environment [11].

In step 4 of IM, the methods and applications are organized in a program plan. This gets round the limitation of the Oinas-Kukkonen model that “does not discuss how individual intervention elements may be varied or integrated into a larger treatment program” [1]. IM has been used to develop eHealth or mHealth interventions [12,13] but is applicable to other intervention types. The BIT model provides tools that are very useful and more specific to the context of eHealth and mHealth interventions (ie, the BIT-Tech aspect of the model).

In sum, both the BIT model and IM address the limitations of existing models. Despite considerable overlap, they each have a unique contribution. Whereas IM stresses the importance of parameters for use, the BIT model focuses on the technical instantiation. The BIT model and IM are complementary, each with their own qualities. For example, when using IM, the BIT-Tech aspect of the BIT model is deemed useful in step 4. On the other hand, when using the BIT model, intervention developers should take the parameters for use from IM into account and report this in the research protocol [14]. In my opinion, the unique contributions of both should be valued during the intervention development process.

References

1. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
2. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for Internet interventions. Ann Behav Med 2009 Aug;38(1):18-27 [FREE Full text] [doi: 10.1007/s12160-009-9133-4] [Medline: 19802647]
3. Fogg BJ. A behavior model for persuasive design. Persuasive’09 :2009.
4. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Communications of the Association for Information Systems 2009;24:485-500.
5. Bartholomew LK, Parcel GS, Kok G, Gottlieb NH, Fernández ME. Planning health promotion programs: An intervention mapping approach. San Francisco, CA: Jossey-Bass; 2011.
6. Kok G, Schaalma H, Ruiter RA, van Empelen P, Brug J. Intervention mapping: protocol for applying health psychology theory to prevention programmes. J Health Psychol 2004 Jan;9(1):85-98. [doi: 10.1177/1359105304038379] [Medline: 14683571]
7. Crutzen R. From eHealth technologies to interventions. J Med Internet Res 2012;14(3):e93 [FREE Full text] [doi: 10.2196/jmir.2050] [Medline: 22738742]
8. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
9. Crutzen R, Cyr D, de Vries NK. Bringing loyalty to e-Health: theory validation using three internet-delivered interventions. J Med Internet Res 2011;13(3):e73 [FREE Full text] [doi: 10.2196/jmir.1837] [Medline: 21946128]
10. Kazdin AE. Behavior Modification in Applied Settings. Long Grove, IL: Waveland Pr Inc; 2008.
11. Forsyth DR. Group Dynamics. Hampshire, UK: Wadsworth Publishing; 2013.
12. Brendryen H, Johansen A, Nesvåg S, Kok G, Duckert F. Constructing a theory- and evidence-based treatment rationale for complex eHealth interventions: Development of an online alcohol intervention using an intervention mapping approach. JMIR Res Protoc 2013;2(1):e6 [FREE Full text] [doi: 10.2196/resprot.2371] [Medline: 23612478]
13. Sassen B, Kok G, Mesters I, Crutzen R, Cremers A, Vanhees L. A web-based intervention for health professionals and patients to decrease cardiovascular risk attributable to physical inactivity: development process. JMIR Res Protoc 2012;1(2):e21 [FREE Full text] [doi: 10.2196/resprot.1804] [Medline: 23612470]
14. Peters GY, de Bruin M, Crutzen R. Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions. Health Psychology Review 2013 Oct 21:1-14. [doi: 10.1080/17437199.2013.848409]

Edited by G Eysenbach; submitted 18.06.14; peer-reviewed by D Mohr; comments to author 29.07.14; revised version received 30.07.14; accepted 04.08.14; published 05.08.14.

Please cite as:
Crutzen R
The Behavioral Intervention Technology Model and Intervention Mapping: The Best of Both Worlds
J Med Internet Res 2014;16(8):e188
URL: http://www.jmir.org/2014/8/e188/
doi:10.2196/jmir.3620
PMID:25095730

Letter to the Editor

Massive Open Online Courses on Health and Medicine: Will They Be Sustainable?

Kieran Walsh[1], FRCPI

BMJ Learning, London, United Kingdom

Corresponding Author:
Kieran Walsh, FRCPI
BMJ Learning
Tavistock Square
London, WC1H 9JR
United Kingdom
Phone: 44 447985755333
Fax: 44 442073836242
Email: kmwalsh@bmjgroup.com

Related Article:

Comment on: http://www.jmir.org/2014/8/e191/

(J Med Internet Res 2014;16(8):e197) doi:10.2196/jmir.3798

In this issue of JMIR, Liyanagunawardena and Williams have provided fascinating insight into the world of massive open online courses (MOOCs) as they have emerged over the past few years [1]. Their findings are clear: there is already a range of MOOCs available and they can be used for a variety of purposes in undergraduate medical education and continuous medical education for medical students, doctors, and health care professionals, with potential in health education amongst the general public.

These are interesting findings and may represent a significant “next step” in the provision of online learning in health care professional education. However, some would say that the past ten years have been as much about the hype of online learning as about the real outcomes that it can actually achieve [2]. This phenomenon is not peculiar to online learning—it happens with virtually all new media when they are initially introduced to education. In past few years however, there has been a shift in thinking about this new delivery mechanism of learning. Exponents of online learning no longer claim that it can do everything or that it will replace face-to-face education, rather, they are starting to think about how it can be used strategically and how its advantages can be adequately harnessed. Such advantages might include its flexibility or increased access to learners. This new and sober atmosphere with regard to online learning in medical education means that it is probably a good time for the medical education community to look at how MOOCs can be harnessed to deliver better education.

Certainly MOOCs satisfy many of the criteria that providers of medical education would like to achieve. They enable increased access, flexibility, and choice to the learner by offering learning at a time and place that suits the learner with substantial amounts of educational content. Another important component is that they are free to the learner. Free access is clearly important to many learners, but the current business model for the provision of MOOCs remains uncertain. Online learning is associated with significant costs, which cannot be ignored in the current economic environment [3]. How long will universities be able to make their content freely available throughout the world without undermining their basic business model, which involves charging learners for their courses? The answer to this question will likely help us draw conclusions as to whether MOOCs are just another passing technology fad or a sustainable long-term solution.

References

1. Liyanagunawardena TR, Williams SA. Massive open online courses on health and medicine: Review. J Med Internet Res 2014;16(8):e191 [FREE Full text] [doi: 10.2196/jmir.3439] [Medline: 25123952]
2. Sandars J. Cost-effective e-learning in medical education. In: Walsh K, editor. Cost Effectiveness in Medical Education (Masterpass S). Oxford: Radcliffe Medical PR; 2010.

3. Cook DA, Triola MM. What is the role of e-learning? Looking past the hype. Med Educ 2014 Sep;48(9):930-937. [doi: 10.1111/medu.12484] [Medline: 25113119]

Edited by G Eysenbach; submitted 19.08.14; peer-reviewed by T Liyanagunawardena, S Williams; accepted 24.08.14; published 25.08.14.

Please cite as:
Walsh K
Massive Open Online Courses on Health and Medicine: Will They Be Sustainable?
J Med Internet Res 2014;16(8):e197
URL: http://www.jmir.org/2014/8/e197/
doi:10.2196/jmir.3798
PMID:25154793

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org