
Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 16 (2014), Issue 10 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

The Impact of an Incentive on the Use of an Online Self-Directed Wellness and Self-Management Program (e217) Judith Hibbard, Jessica Greene.	3
Development and Preliminary Evaluation of an Internet-Based Healthy Eating Program: Randomized Controlled Trial (e231) Katy Tapper, Gabriela Jiga-Boy, Gregory Maio, Geoffrey Haddock, Michael Lewis.	12
Posttreatment Attrition and Its Predictors, Attrition Bias, and Treatment Efficacy of the Anxiety Online Programs (e232) Ali AL-Asadi, Britt Klein, Denny Meyer.	28
Understanding Usage of a Hybrid Website and Smartphone App for Weight Management: A Mixed-Methods Study (e201) Leanne Morrison, Charlie Hargood, Sharon Lin, Laura Dennison, Judith Joseph, Stephanie Hughes, Danius Michaelides, Derek Johnston, Marie Johnston, Susan Michie, Paul Little, Peter Smith, Mark Weal, Lucy Yardley.	43
Diabetes Patients' Experiences With the Implementation of Insulin Therapy and Their Perceptions of Computer-Assisted Self-Management Systems for Insulin Therapy (e235) Airin Simon, Wouter Gude, Frits Holleman, Joost Hoekstra, Niels Peek.	63
Evaluating the Interactive Web-Based Program, Activate Your Heart, for Cardiac Rehabilitation Patients: A Pilot Study (e242) Christopher Brough, Sally Boyce, Linzy Houchen-Wolloff, Louise Sewell, Sally Singh.	77
A Comparison Between Phone-Based Psychotherapy With and Without Text Messaging Support In Between Sessions for Crisis Patients (e219) Gareth Furber, Gabrielle Jones, David Healey, Niranjana Bidargaddi.	88
Use of Mobile Phone Text Message Reminders in Health Care Services: A Narrative Literature Review (e222) Kati Kannisto, Marita Koivunen, Maritta Välimäki.	100
Beyond Traditional Advertisements: Leveraging Facebook's Social Structures for Research Recruitment (e243) Rupa Valdez, Thomas Guterbock, Morgan Thompson, Jeremiah Reilly, Hannah Menefee, Maria Bennici, Ishan Williams, Deborah Rexrode.	1

Tweet Content Related to Sexually Transmitted Diseases: No Joking Matter (e228)	
Elia Gabarron, J Serrano, Rolf Wynn, Annie Lau.	134
Tweeting for and Against Public Health Policy: Response to the Chicago Department of Public Health's Electronic Cigarette Twitter Campaign (e238)	
Jenine Harris, Sarah Moreland-Russell, Bechara Choucair, Raed Mansour, Mackenzie Staub, Kendall Simmons.	144
Rapid Grading of Fundus Photographs for Diabetic Retinopathy Using Crowdsourcing (e233)	
Christopher Brady, Andrea Villanti, Jennifer Pearson, Thomas Kirchner, Omesh Gupta, Chirag Shah.	157
Identifying Problematic Internet Users: Development and Validation of the Internet Motive Questionnaire for Adolescents (IMQ-A) (e230)	
Christina Bischof-Kastner, Emmanuel Kuntsche, Jörg Wolstein.	167
Comorbidity Structure of Psychological Disorders in the Online e-PASS Data as Predictors of Psychosocial Adjustment Measures: Psychological Distress, Adequate Social Support, Self-Confidence, Quality of Life, and Suicidal Ideation (e248)	
Ali AL-Asadi, Britt Klein, Denny Meyer.	179
What Are Health-Related Users Tweeting? A Qualitative Content Analysis of Health-Related Users and Their Messages on Twitter (e237)	
Joy Lee, Matthew DeCamp, Mark Dredze, Margaret Chisolm, Zackary Berger.	194
A Case Study of the New York City 2012-2013 Influenza Season With Daily Geocoded Twitter Data From Temporal and Spatiotemporal Perspectives (e236)	
Ruchit Nagar, Qingyu Yuan, Clark Freifeld, Mauricio Santillana, Aaron Nojima, Rumi Chunara, John Brownstein.	203
Evaluating the Process of Online Health Information Searching: A Qualitative Approach to Exploring Consumer Perspectives (e224)	
Alexander Fiksdal, Ashok Kumbamu, Ashutosh Jadhav, Cristian Cocos, Laurie Nelsen, Jyotishman Pathak, Jennifer McCormick.	218
Tablet, Web-Based, or Paper Questionnaires for Measuring Anxiety in Patients Suspected of Breast Cancer: Patients' Preferences and Quality of Collected Data (e239)	
Maarten Barentsz, Hester Wessels, Paul van Diest, Ruud Pijnappel, Cees Haaring, Carmen van der Pol, Arjen Witkamp, Maurice van den Bosch, Helena Verkooijen.	227
Digital Inequalities of Family Life Information Seeking and Family Well-Being Among Chinese Adults in Hong Kong: A Population Survey (e227)	
Man Wang, Xin Wang, Kasisomayajula Viswanath, Alice Wan, Tai Lam, Sophia Chan.	236
Supporting Health Care Professionals to Improve the Processes of Shared Decision Making and Self-Management in a Web-Based Intervention: Randomized Controlled Trial (e211)	
Barbara Sassen, Gerjo Kok, Jan Schepers, Luc Vanhees.	247
Patients' Continuing Use of an Online Health Record: A Quantitative Evaluation of 14,000 Patient Years of Access Data (e241)	
Richard Phelps, Joanne Taylor, Keith Simpson, Jasmine Samuel, A Turner.	262
Promoting Health in Virtual Worlds: Lessons From Second Life (e229)	
Reima Suomi, Matti Mäntymäki, Sari Söderlund.	277
Automatic Evidence Retrieval for Systematic Reviews (e223)	
Miew Choong, Filippo Galgani, Adam Dunn, Guy Tsafnat.	286

Original Paper

The Impact of an Incentive on the Use of an Online Self-Directed Wellness and Self-Management Program

Judith H Hibbard¹, DrPH; Jessica Greene², PhD

¹University of Oregon-1209, Health Policy Research Group, Eugene, OR, United States

²George Washington University, School of Nursing, Washington, DC, United States

Corresponding Author:

Judith H Hibbard, DrPH
University of Oregon-1209
Health Policy Research Group
119 Hendricks Hall
Eugene, OR, 97403-1209
United States
Phone: 1 541 346 3364
Fax: 1 541 346 2040
Email: jhibbard@uoregon.edu

Abstract

Background: Those who pay for health care are increasingly looking for strategies to influence individuals to take a more active role in managing their health. Incenting health plan members and/or employees to participate in wellness programs is a widely used approach.

Objective: In this study, we examine financial incentives to health plan members to participate in an online self-management/wellness program—US \$20 for completing the patient activation measure (PAM) and an additional US \$40 for completing 8 learning modules. We examined whether the characteristics of plan members differed by the degree to which they responded to the incentives. Further, we examined whether participation in the wellness program was associated with improvements in PAM scores and changes in health care utilization.

Methods: This retrospective study compared demographic characteristics and change in PAM scores and health utilization for 144,625 health plan members in 2011. Four groups were compared: (1) those who were offered the incentives but chose not to participate (n=128,634), (2) those who received the initial incentive (PAM only) but did not complete 8 topics (n=7099), (3) those who received both incentives (completing 8 topics but no more) (n=2693), and (4) those who received both incentives and continued using the online program beyond what was required by the incentives (n=6249).

Results: The vast majority of health plan members did not participate in the program (88.91%, 128,634/144,675). Of those who participated, only 7099 of 16,041 (44.25%) completed the PAM for the first incentive, 2693 (16.79%) completed 8 topics for the second incentive, and 6249 (38.96%) received both incentives and continued using the program beyond the incentive requirements. Nonparticipants were more likely to be men and to have lower health risk scores on average than the other three groups of participants ($P<.001$). In multivariate regression models, those who used the online program (8 topics or beyond) increased their PAM score by approximately 1 point more than those who only took the PAM and did not use the wellness program ($P<.03$). In addition, emergency department visits were lower for all groups who responded to any level of the incentive as compared to those who did not ($P<.01$). No differences were found in other types of utilization.

Conclusions: The incentive was not sufficient to spark most health plan members to use the wellness program. However, the fact that many program participants went beyond the incentive in their use of the online wellness program suggests that the users of the online program found value in using it, and it was their own internal motivation that stimulated this additional use. Providing an incentive for program participation may be an effective pathway for working with less activated patients, particularly if the program is tailored to the needs of the less activated.

(*J Med Internet Res* 2014;16(10):e217) doi:[10.2196/jmir.3239](https://doi.org/10.2196/jmir.3239)

KEYWORDS

empowerment; health promotion; Internet; incentives; patient activation

Introduction

Theory and Prior Work

Patients are key in determining their health care outcomes. They are the ones that carry out the day-to-day care management tasks and decide whether or not to make the necessary lifestyle adjustments to improve their health. Without patients' participation, even with best practices on the part of health care providers, it is very difficult to achieve optimal health outcomes [1]. To control costs and improve outcomes, insurance companies, government, and employers are increasingly looking for strategies to influence individuals to take a more active role in managing their health.

Incenting health plan members and/or employees to take Health Risk Appraisals (HRAs) is a widely used approach. The HRAs typically include questions about health behaviors and provide the individual feedback with a list of behavioral changes they could make to improve their health. A 2013 study estimated that 55% of large employers offered an HRA to employees and 54% of those provided an incentive to complete it [2]. That was up from 49% offering the HRA five years ago in 2008 and 33% incenting it. The use of HRAs is based on the assumption that raising awareness of health risks and needed behavioral changes will be sufficient to stimulate change, an assumption that is often not borne out [3]. Now the trend has expanded to include incentives that are tied to participation in a wellness program. For example, the 2013 Health Research and Educational Trust (HRET) survey found that 99% of large firms offer wellness programs and 36% incent participation in them [2]. Towers Watson Staying@Work Survey Report found that in 2011, half of employees surveyed were offered financial rewards for participation in a health program [4]. Under the Affordable Care Act (ACA) in the United States, new rules will allow employers to increase the use of incentives for wellness programs [5].

Studies indicate that incentives are most effective at achieving behavior change that requires a single activity, such as receiving a vaccination or a screening, as compared to actions that require ongoing engagement [6,7]. There is concern that when health plans or insurers offer incentives, discrimination or disadvantages do not accrue to those who choose not to participate. Of particular concern is when a specific health behavior is the focus, for example smoking, that the individual who opts out of participation is not penalized at work or in their access to insurance [8].

However, there are fewer concerns about the use of incentives to complete an HRA. There is evidence that incentives do yield greater participation in HRAs [9,10]. The degree to which incentives stimulate meaningful engagement in longer-term health programs is less clear. There is concern that financial incentives may supplant or undermine internal motivations. The worry is that the incentive may undermine a sense of personal responsibility for health and an individual's intrinsic motivations to promote their own health [11].

Might an incentive to participate in a wellness/self-management program undercut internal motivation? One way to examine this question is to assess the impact of an incentive on program

participation and on changes in patient activation. Patient activation is defined as having the knowledge, skill, and confidence to manage one's own health and health care. Individuals who are more activated, are typically more motivated to effectively manage their health. While there are alternative ways to define activation or engagement, there is only one validated measure, the 13-item Patient Activation Measure (PAM). Multiple studies show that people who score higher on PAM are more motivated to improve their health. The PAM score is predictive of most health behaviors, many clinical indicators, and the use of costly health care services [12,13]. If PAM scores were to improve as a result of incented program participation, this would indicate that the incentive did not undercut internal motivation.

Most of the research on the efficacy of health program participation is based on programs offered at the worksite or programs offered in the community. Because wellness and self-management support programs are now also being offered as online programs, they are no longer limited to these venues. Studies indicate that there is some positive impact of self-directed online self-management programs. For example, Solomon and colleagues, evaluating the impact of an online program among chronic disease patients, found that the program increased patient activation an average of 4 points (on a 100-point scale) more than did the control group [14]. Similarly, Lorig and colleagues found that an online self-management program improved clinical outcomes and patient activation [15]. However, there appears to be no evidence on the value of offering incentives to participate in these self-directed online programs.

Research Questions

In this study, we examine the impact of providing financial incentives to commercially insured health plan members to participate in an online self-directed self-management/wellness program. We assess the impact of a US \$20 financial incentive to take a PAM assessment and a second US \$40 incentive to complete 8 user-selected learning modules or topics in the online program and to set 8 behavioral goals.

Our specific research questions were: (1) What are the characteristics of those who respond to the two levels of financial incentives and those who do not respond at all to the incentive?; (2) What are the characteristics of those who participate beyond the financial incentive in using the wellness/self-management program? How much further beyond the incentive does their participation go?; and (3) What is the impact of participating in the online program on patient activation and health care utilization?

Methods

Incentive

In 2011, commercially insured members of a large health plan were offered a US \$20 incentive to go online to take the PAM questions. They were given a further US \$40 incentive to complete 8 topics and set behavioral goals in the online self-directed wellness/self-management program. The participant could choose the topic areas and behavioral goals they wanted

to focus on. Thus there are four groups in the study: (1) nonparticipants in the online program, (2) those who only completed the PAM and received the first incentive, (3) those who completed the PAM and 8 topics and received the first and second incentive, and (4) those who received both incentives and continued using the online program beyond the required number of topics.

Online Wellness Program

Participants were incented to go to an online program that assessed their level of patient activation and to participate in a self-directed online wellness/self-management program. All the topics and suggested goals that users are able to select from in the online program are tailored to the users' level of activation. The program tailors to the 4 levels of activation, from low (level 1) to high (level 4). Users can choose what they want to focus on, from disease specific self-management to wellness activities. The lowest activation level topics start at a very foundational level. The program breaks information down into small bits and suggests action steps or goals that less activated individuals are likely to succeed at, but may not be clinically meaningful at this point (eg, cut out fast food lunches 2 days a week). The approach is based on the theories of behavioral activation that show motivation will follow action [16]. The strategy is to start the individual acting, experiencing small successes, and those successes lead to greater motivation and confidence. Level 4 topics assume both a higher level of knowledge and skill and encourage users to set behavioral goals that, for example, help them stay on track with behaviors they have already adopted. Users focus on the topics they are interested in and set behavioral goals related to the topics they pursue. The program provides insights and strategies for overcoming barriers to meeting their goals.

Design and Analysis

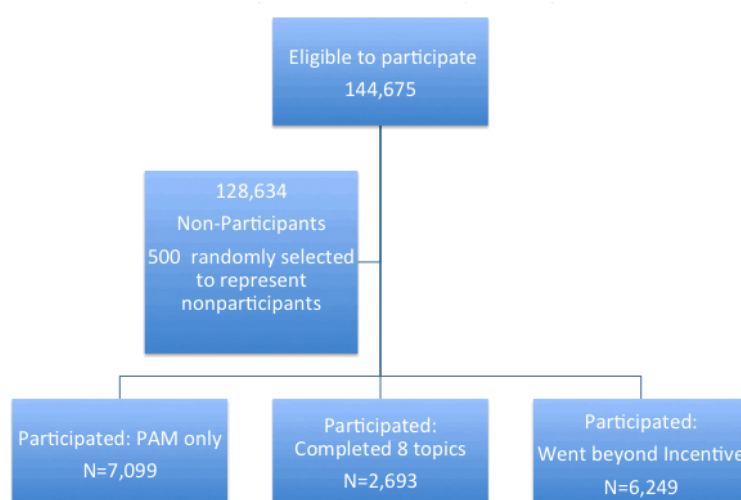
This retrospective study compared demographic characteristics and change in PAM and health utilization for four groups of health plan members in 2011: (1) those who were offered the incentives but chose not to participate in the online wellness program, (2) those who received the initial incentive (PAM only) but did not complete 8 topics and goals, (3) those who received both incentives (completing 8 topics and 8 goals but no more), and (4) those who received both incentives and continued using the online wellness program beyond what was required by the incentives. The study protocol (designated as 12012011.001), was approved by the internal review boards (IRB) of the University of Oregon and George Washington University.

Study Population

The total population eligible for participation was 144,675 health plan members who were adults and had consistent coverage over the study period. Given that 128,634 people declined to respond to the incentive (88.91% of the eligible population did not use the online wellness program), we selected a random sample of 500 of these nonparticipating plan members to represent the nonparticipant group. The other 3 groups included all of those who chose to respond to the incentives and individually decided what level of incentive they would pursue (Figure 1). Observations are weighted in all analyses using the inverse probability of selection in order to reflect the percent of the total population they represent.

The study population excluded those plan members who were recruited for a special program that involved telephonic coaching because of serious health issues. We excluded this group whether they participated in the telephonic coaching or not. These excluded patients were sicker with more chronic conditions and higher levels of hospitalization and emergency department use than the larger population of plan members.

Figure 1. Study design showing the study groups.



Measures

The independent variable in this study was the extent to which the eligible health plan member used the online wellness

program (no participation, PAM only/low incentive, completed 8 topics/both incentives, went beyond incentive requirements).

Dependent variables included the PAM and health care utilization measures. The PAM consists of 13 items that form

an interval level, unidimensional, Guttman-like scale. The measure has been shown to have strong psychometric properties. The PAM items are statements on confidence, beliefs, knowledge, and skills about managing one's health, which respondents can answer with degrees of agreement or disagreement (eg, I know how to prevent problems with my health; I am confident that I can tell a doctor my concerns, even when he or she does not ask). The measure is scored on a theoretical 0-100 scale with most patients falling in the 35-95 point range. Four levels of activation have been previously identified, reflecting a developmental progression from passive receipt of care toward greater activation [17,18].

Utilization data were derived from claims data. Primary care, specialty care, emergency department (ED) use, and hospital stays were the utilization variables assessed. For primary and specialty care visits, we examined the actual number of visits in the pre and post period. For the ED and hospitalizations, we examined whether or not there were any visits.

In our analyses we also examined and controlled for demographic and health variables. Average income for the member's zip code is used as a proxy for member's income in the analysis. The Adjusted Clinical Groups (ACG) score is a measure of illness severity and is included in the analysis to control for health status differences in the study groups.

Analytic Approach

We first examined the demographic and health characteristics of the four groups, using baseline data from the last 6 months of 2010, using chi-square and analysis of variance (ANOVA) tests. Additionally, we used descriptive statistics to examine how much use of the online wellness program each of the four groups engaged in.

Then, we examined bivariate relationship between the changes from pre- to post- in the outcome measures (PAM scores and utilization) using follow-up data collected from the first 6 months of 2012. Finally, we developed multivariate analyses to examine the relationships between group participation level and the post-period outcomes, controlling for the pre-period outcome value, demographic characteristics, and ACG risk score. The sample size for Patient activation in the follow-up is reduced to 46% of the original sample size because not all members completed the pam again in 2012. All analyses weight the observations by the inverse probability of selection and adjust the standard errors accordingly.

Results

Characteristics of Those Who Responded to Each Level of the Incentive

Table 1 shows that the vast majority of study participants did not use the online wellness program at all (88.91%,

128,634/144,675). Of the 16,041 people who did use it, 7099 (44.25%) people participated only enough to receive the initial incentive, 2693 (16.79%) received the second higher incentive, and 6249 (38.96%) went beyond the incentive requirements.

The people who participated in the four different levels of the online wellness program had different demographic characteristics (Table 1). The group that did not participate in the online wellness program had a higher percent of males than the groups that did respond, and women were more likely to comprehensively use the online program. There were also age differences in who and how much members responded to the incentive. Those that responded to the higher incentive only were slightly younger than those in the other groups. The average income (based on zip code) was not different for the four groups. People whose online program use went beyond the incentive requirements had a slightly higher baseline PAM score than did those who did only what the incentives required (there were no PAM scores on the group that did not participate in the online program at all).

There were also differences in the groups in terms of their health status. Those who responded to the incentive or went beyond the incentive had slightly higher ACG risk scores than those not responding to the incentive. However, those not responding to any of the incentives were more likely to have diabetes and hypertension than those who responded to the incentive or went beyond the incentive.

It is interesting to note that those members who went beyond the incentive in their use of the online program, had usage considerably beyond the requirements of the incentive, viewing on average over 20 topics (Table 2); 69.88% (6249/8942) of those who got the maximum incentive, went beyond what the incentive required.

We also looked at baseline PAM level as a predictor of how members responded to the incentive. Table 3 shows percent of members at each level of activation at baseline and their distribution over the three groups responding to the incentive (no PAM scores were available for those who did not respond to the incentive so they are not included in this analysis). Members at the lowest level of activation at baseline were slightly more likely to take the minimum incentive (PAM only), than those at the higher levels of activation. However, even among the least activated, 32.0% (231/723) went beyond the incentive in their use of the online program, while 40.50% (2828/6983) of the highest activated went beyond the incentive.

Table 1. Demographic and health characteristics of the study sample, by program participation level.

Demographic/characteristic	Total sample ^a (n=16,493)	Subsample of non-participants ^a (n=452)	PAM ^b only: Low incentive participation (n=7099)	Completed 8 topics: High incentive (n=2693)	Went beyond incentive requirements (n=6249)	P value
Weighted percent of population	-	14,609 (88.58%)	870 (5.27%)	305 (1.85%)	709 (4.30%)	
Gender, n (%)						
Male	8881 (53.85)	8048 (55.09)	366 (42.06)	153 (50.16)	314 (44.29) ^c	<.001
Female	7612 (46.15)	6561 (44.91)	504 (57.93)	152 (49.84)	395 (55.71) ^c	
Age, mean (SD)	41.3 (13.4)	41.2 (13.5)	42.8 (12.2)	39.2 (11.8)	41.9 (11.6) ^c	<.001
Income of zip code, mean (SD)	23,967.34 (7131.31)	23,947.40 (7164.20)	24,318.16 (6899.29)	24,137.45 (7199.90)	23,874.46 (6846.15)	.48
PAM score, mean (SD)	N/A	N/A	70.3 (15.3)	70.5 (15.0)	71.7 (15.3) ^c	<.001
Health Risk Score, mean (SD)	2.0 (1.3)	1.9 (1.3)	2.1 (1.2)	2.1 (1.2)	2.2 (1.1) ^c	<.001
Chronic conditions, n (%)						
Depression	2156 (13.07)	1939 (13.27)	103 (11.84)	30 (9.84)	83 (11.71)	.24
Diabetes	519 (3.15)	485 (3.32)	15 (1.72)	5 (1.64)	14 (1.97) ^a	.02
Hypertension	2307 (13.99)	2101 (14.38)	93 (10.69)	28 (9.18)	85 (11.99) ^a	.02

^aSince we examined a random sample of people who never participated in online coaching, the descriptive statistics are weighted using the inverse probability of selection for the no coaching group, and the standard errors are weighted accordingly. The numbers presented in the table for percentages reflect the weighted numbers.

^bPAM: patient activation measure

^cChi-square statistical tests were used for categorical demographic and health variables (gender and having specific chronic conditions), while ANOVAs were used for continuous variables (age, mean income of zip code, health risk score).

Table 2. Incentive response and use of online wellness/ self-management program.

Program use	PAM ^a only: Low incentive mean (SD)	Completed 8 topics: High incentive mean (SD)	Went beyond incentive requirements mean (SD)	ANOVA P value
Number of health topics accessed	1.2 (2.2)	8.6 (1.0)	20.2 (18.4)	<.001
Number of health topics completed	0.9 (1.9)	8.0 (0.0)	19.4 (18.2)	<.001
Number of months of use	1.2 (0.6)	1.4 (0.8)	1.7 (1.1)	<.001

^aPAM: patient activation measure

Table 3. Patient activation level and response to incentive.^a

Response	PAM ^b Level 1 (n=723)	PAM Level 2 (n=991)	PAM Level 3 (n=7342)	PAM Level 4 (n=6983)
	n (%)	n (%)	n (%)	n (%)
PAM only: Low Incentive	388 (53.67)	445 (44.90)	3237 (44.09)	3027 (43.35)
Completed 8 topics: High Incentive	106 (14.66)	179 (18.06)	1282 (17.46)	1126 (16.12)
Went beyond incentive requirements	229 (31.67)	367 (37.03)	2823 (38.45)	2830 (40.53)

^aChi-square P value <.001

^bPAM: patient activation measure

Impact of Intervention on Patient Activation and Utilization

Table 4 shows the bivariate relationships between the program participation level and the pre, post, and change in outcome variables. The results indicate that there is a significant difference in PAM scores in the pre and post periods across the three groups who used the online program (we did not have PAM data on those who did not participate); however, the difference in the increase from baseline to follow-up is not significant.

There were no significant differences in the groups in primary care, specialty care, or hospital visits in either the baseline or post period. However, there were differences in the groups in the post period for ED visits, with highest rates for nonparticipants. It is notable that ED visits increased for all groups from the baseline to the post period, but the change was not significantly different among the groups.

Table 5 shows the multivariate regression results, which control for baseline outcome value, age, gender, income, and baseline ACG risk score. In this analysis, patient activation scores at follow-up are significantly greater (by approximately 1 point) for those who used the online wellness program as compared to those who only took the PAM. In addition, use of the ED is significantly lower for all groups who responded to any level of the incentive as compared to those who did not participate. No differences were found in other types of utilization.

Table 6 shows how much patient activation scores changed among those who made any use of the online program. The changes in scores are shown within baseline PAM levels. At baseline, those who scored in the lower two levels of patient activation increased their PAM score the most following participating in the online program. Members who at baseline scored in the lowest level of activation gained an average of 21 points (on a 0-100 scale).

Table 4. Bivariate relationships between utilization, cost, and participation level.

	No participation	PAM ^c only: Low incentive participation	Completed 8 topics: High incentive	Went beyond incentive requirements	<i>P</i> value
Patient activation score^a, mean (SD)					
Pre-period	n/a	70.7 (14.6)	71.2 (14.7)	72.1 (15.3)	.002
Post-period	n/a	72.4 (14.6)	73.6 (15.1)	74.1 (16.2)	<.001
Change	n/a	1.7 (4.8)	2.5 (15.5)	2.0 (16.3)	.33
Primary care visits, mean (SD)					
Pre-period	2.4 (5.4)	2.4 (4.8)	2.4 (4.6)	2.6 (5.0)	.47
Post-period	3.1 (4.9)	3.5 (5.3)	3.1 (5.6)	3.4 (5.9)	.06
Change	0.7 (6.8)	1.0 (6.5)	0.7 (6.9)	0.8 (7.1)	.50
Specialty care visits, mean (SD)					
Pre-period	2.9 (7.2)	2.8 (6.5)	2.8 (5.9)	2.9 (6.4)	.86
Post-period	3.7 (9.1)	4.1 (8.0)	3.6 (7.3)	3.9 (7.9)	.10
Change	0.8 (9.5)	1.3 (8.7)	0.8 (8.3)	1.0 (8.9)	.20
Emergency department visits, n (%) with a visit^b					
Pre-period	635 (4.5)	27 (3.0)	9 (3.1)	24 (3.4)	.41
Post-period	1103 (7.8)	33 (3.7)	12 (4.0)	31 (4.3)	.01
Change	3.3 (0.3)	0.7 (0.3)	0.9 (0.3)	0.9 (0.3)	.40
Hospitalizations, n (%) with a visit^b					
Pre-period	226 (1.6)	10 (1.1)	3 (1.0)	8 (1.2)	.81
Post-period	291 (2.0)	17 (2.0)	7 (2.2)	16 (2.2)	.83
Change	0.4 (0.2)	0.8 (0.2)	1.2 (0.2)	1.1 (0.2)	.74

^aThe sample size for PAM was n=2222 for PAM only, n=1522 completed 8 topics, and n=3894 for going beyond the incentive requirements.

^bThe numbers presented in the table for percentages reflect the weighted numbers.

^cPAM: patient activation measure

Table 5. Key regression coefficients from models examining follow-up PAM^a and health care utilization, by participation status.^b

Participation	PAM		Primary care visits		Specialty care visits		Any ED ^c visits		Any hospitalizations	
	Coeff	P value	Coeff	P value	Coeff	P value	Coeff	P value	Coeff	P value
No Participation	n/a		--		--		--		--	
PAM only: Low incentive participation	--		.01	.97	.16	.73	-.05	.005	.00	.62
Completed 8 topics: High incentive only	.94	.03	-.14	.59	-.21	.65	-.05	.007	.01	.26
Went beyond incentive requirements	.99	.005	-.05	.86	-.04	.93	-.04	.01	-.01	.28

^aPAM: patient activation measure

^bModels control for baseline value of the dependent variable, gender, age, terciles of mean income of zip code, and baseline Adjusted Clinical Groups (ACG) risk score.

^cED: emergency department

Table 6. PAM^a score changes among those responding to incentive.

	PAM Level 1 at baseline	PAM Level 2 at baseline	PAM Level 3 at baseline	PAM Level 4 at baseline	P value
	mean (SD)	mean (SD)	mean (SD)	mean (SD)	
Any use of online program	21.0 (22.2)	11.5 (13.1)	5.2 (13.0)	-4.0 (15.1)	<.001

^aPAM: patient activation measure

Discussion

Principal Findings

Only about 11% of those offered an incentive to participate in an online wellness program participated in it. Those who participated in any level of the incentive were more likely to be female, slightly older, and less likely to have diabetes, hypertension, and depression than those who did not respond. We did not have PAM data on the group who did not respond to the incentive; however, among those who did respond, those with slightly higher scores at baseline were more likely to go beyond the incentive in using the online program. That is to say, the incentive seemed to be most effective in stimulating participation among those who were slightly more activated and in better health. This is consistent with a prior study indicating that it is the more activated who are more likely to participate in offered health programs [19]. It is also consistent with findings that suggest that incentives are more effective when the incentive does not require difficult or sustained behavior change [7].

Contrary to concerns raised by some investigators, the findings do not indicate that the incentive to participate in the online program undermined or supplanted internal motivations to manage one's own health [11]. For those who used the online wellness program, 40% went beyond what the incentive required, and 70% of those who took the maximum incentive, continued to use the program beyond the incentive. This suggests that the users of the online program found value in using it, and it was their own internal motivation that stimulated this additional use. Further, it appears that participation in the online program increased PAM scores, particularly among those who began at the lower two levels of activation. The findings also show a lesser increase in emergency department use for

those who participated in any way in the program compared to those who did not use the online wellness program. Given that the sickest and costliest patients were excluded from the analysis (those invited to telephonic coaching were excluded), it is somewhat surprising that this type of "self-directed" intervention would impact utilization among those who are moderate to low health care utilizers.

The findings break new ground in this research arena, in that most evaluations of wellness programs and incentives do not measure the impact of the program on the individual's knowledge, skill, and confidence in managing their health (activation). By including activation in the analysis, the findings help to elucidate who these programs are reaching (eg, are they reaching the more or less activated?) and who they are helping.

Limitations

At the same time, the findings are limited by the fact that we did not have PAM data on those who chose not to participate in the incentive program, nor on participants who did not complete a PAM in 2012 as well as 2011. They are also limited by the design of the study. Because participants were not randomly assigned to the four study groups, we cannot rule out that the findings are the result of some other, unmeasured, factor. Finally, because the study was conducted within one health plan in one geographic region and only included insured adults, the generalizability of the findings are limited.

The findings indicate that participation in the program resulted in a small decline in the PAM scores of those who, at baseline, were in the highest level of activation. It may be that the intervention was not as useful or appropriate for these participants. However, the fact that program participation appeared to stimulate increased activation in those who were the least activated at baseline is a promising finding. These

findings are consistent with the findings from other intervention studies that show that the least activated gain the most from an intervention [14,15,20]. Research shows that increases in activation are related to improvements in health behaviors, clinical indicators, and reductions in costly health care utilization [1]. As we move into an era where organizations are held accountable for costs and outcomes, finding ways to engage the least activated is becoming a priority.

Conclusions

While the incentive was most effective in recruiting higher activated members, less activated members also responded to

the incentive. Providing an incentive for program participation may be an effective pathway for working with less activated patients, particularly if the program is tailored to the needs of the less activated. However, it is important to note that the vast majority of health plan members did not respond to the incentive at all.

With increased costs and an aging population base, strategies to effectively manage population health is a global concern. This study adds to our understanding of who is likely to respond to incentives and how participation in the programs may influence behaviors and health care utilization.

Acknowledgments

The authors wish to acknowledge the funding provided by Medica Research Institute and the assistance of the Medica Research Institute staff. A special thanks to Ravi Narayanan, who provided assistance in securing the data, data interpretation, and in interpretation of elements of the program.

Conflicts of Interest

Judith Hibbard is a consultant to and an equity stakeholder in Insignia Health. Jessica Greene has no conflicts of interest to report.

References

1. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. *J Gen Intern Med* 2012 May;27(5):520-526 [FREE Full text] [doi: [10.1007/s11606-011-1931-2](https://doi.org/10.1007/s11606-011-1931-2)] [Medline: [22127797](https://pubmed.ncbi.nlm.nih.gov/22127797/)]
2. Claxton G, Rae M, Panchal N, Damico A, Whitmore H, Bostick N, et al. Health benefits in 2013: moderate premium increases in employer-sponsored plans. *Health Aff (Millwood)* 2013 Sep;32(9):1667-1676. [doi: [10.1377/hlthaff.2013.0644](https://doi.org/10.1377/hlthaff.2013.0644)] [Medline: [23962411](https://pubmed.ncbi.nlm.nih.gov/23962411/)]
3. Huskamp HA, Rosenthal MB. Health risk appraisals: how much do they influence employees' health behavior? *Health Aff (Millwood)* 2009 Oct;28(5):1532-1540 [FREE Full text] [doi: [10.1377/hlthaff.28.5.1532](https://doi.org/10.1377/hlthaff.28.5.1532)] [Medline: [19738273](https://pubmed.ncbi.nlm.nih.gov/19738273/)]
4. The Advisory Board Company. The Daily Briefing. 2012 Aug 13. Employers increasingly incent wellness accountability, not just participation URL: <http://www.advisory.com/daily-briefing/2012/08/13/employers-increasingly-incent-wellness-accountability-not-just-participation> [accessed 2014-09-16] [WebCite Cache ID [6SdHxH0IO](https://www.webcitation.org/6SdHxH0IO)]
5. The Kaiser Family Foundation. Kaiser Health News. 2013 May 30. New Rule Allows Employers to Increase the Use of Wellness Incentives URL: <http://www.kaiserhealthnews.org/daily-reports/2013/may/30/health-reform-wellness-rule.aspx> [accessed 2014-09-16] [WebCite Cache ID [6SdI7K6iZ](https://www.webcitation.org/6SdI7K6iZ)]
6. Kane RL, Johnson PE, Town RJ, Butler M. A structured review of the effect of economic incentives on consumers' preventive behavior. *Am J Prev Med* 2004 Nov;27(4):327-352. [doi: [10.1016/j.amepre.2004.07.002](https://doi.org/10.1016/j.amepre.2004.07.002)] [Medline: [15488364](https://pubmed.ncbi.nlm.nih.gov/15488364/)]
7. Sutherland K, Christianson JB, Leatherman S. Impact of targeted financial incentives on personal health behavior: a review of the literature. *Med Care Res Rev* 2008 Dec;65(6 Suppl):36S-78S. [doi: [10.1177/1077558708324235](https://doi.org/10.1177/1077558708324235)] [Medline: [19015378](https://pubmed.ncbi.nlm.nih.gov/19015378/)]
8. Dixon A. LIVESTRONG. 2013 Aug 16. Ethical Implications of Wellness Programs URL: <http://www.livestrong.com/article/502335-ethical-implications-of-wellness-programs/> [accessed 2014-09-16] [WebCite Cache ID [6SdIn2HY0](https://www.webcitation.org/6SdIn2HY0)]
9. Seaverson EL, Grossmeier J, Miller TM, Anderson DR. The role of incentive design, incentive value, communications strategy, and worksite culture on health risk assessment participation. *Am J Health Promot* 2009;23(5):343-352. [doi: [10.4278/ajhp.08041134](https://doi.org/10.4278/ajhp.08041134)] [Medline: [19445438](https://pubmed.ncbi.nlm.nih.gov/19445438/)]
10. Taitel MS, Haufle V, Heck D, Loeppeke R, Fetterolf D. Incentives and other factors associated with employee participation in health risk assessments. *J Occup Environ Med* 2008 Aug;50(8):863-872. [doi: [10.1097/JOM.0b013e3181845fe2](https://doi.org/10.1097/JOM.0b013e3181845fe2)] [Medline: [18695446](https://pubmed.ncbi.nlm.nih.gov/18695446/)]
11. Lunze K, Paasche-Orlow MK. Financial incentives for healthy behavior: ethical safeguards for behavioral economics. *Am J Prev Med* 2013 Jun;44(6):659-665. [doi: [10.1016/j.amepre.2013.01.035](https://doi.org/10.1016/j.amepre.2013.01.035)] [Medline: [23683984](https://pubmed.ncbi.nlm.nih.gov/23683984/)]
12. Hibbard JH, Greene J, Overton V. Patients with lower activation associated with higher costs; delivery systems should know their patients' 'scores'. *Health Aff (Millwood)* 2013 Feb;32(2):216-222. [doi: [10.1377/hlthaff.2012.1064](https://doi.org/10.1377/hlthaff.2012.1064)] [Medline: [23381513](https://pubmed.ncbi.nlm.nih.gov/23381513/)]
13. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. *Health Aff (Millwood)* 2013 Feb;32(2):207-214. [doi: [10.1377/hlthaff.2012.1061](https://doi.org/10.1377/hlthaff.2012.1061)] [Medline: [23381511](https://pubmed.ncbi.nlm.nih.gov/23381511/)]

14. 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](https://doi.org/10.2196/jmir.1924)] [Medline: [22353433](https://pubmed.ncbi.nlm.nih.gov/22353433/)]
15. Lorig K, Alvarez S. Re: Community-based diabetes education for Latinos. *Diabetes Educ* 2011 Feb;37(1):128. [doi: [10.1177/0145721710393089](https://doi.org/10.1177/0145721710393089)] [Medline: [21292624](https://pubmed.ncbi.nlm.nih.gov/21292624/)]
16. Sturmey P. Behavioral activation is an evidence-based treatment for depression. *Behav Modif* 2009 Nov;33(6):818-829. [doi: [10.1177/0145445509350094](https://doi.org/10.1177/0145445509350094)] [Medline: [19933444](https://pubmed.ncbi.nlm.nih.gov/19933444/)]
17. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: [10.1111/j.1475-6773.2005.00438.x](https://doi.org/10.1111/j.1475-6773.2005.00438.x)] [Medline: [16336556](https://pubmed.ncbi.nlm.nih.gov/16336556/)]
18. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res* 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: [10.1111/j.1475-6773.2004.00269.x](https://doi.org/10.1111/j.1475-6773.2004.00269.x)] [Medline: [15230939](https://pubmed.ncbi.nlm.nih.gov/15230939/)]
19. Hibbard JH, Greene J. Who are we reaching through the patient portal: engaging the already engaged? *International Journal of Person Centered Medicine* 2011 Dec;1(4):788-793. [doi: [10.5750/ijpcm.v1i4.152](https://doi.org/10.5750/ijpcm.v1i4.152)]
20. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VB, et al. Online diabetes self-management program: a randomized study. *Diabetes Care* 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: [10.2337/dc09-2153](https://doi.org/10.2337/dc09-2153)] [Medline: [20299481](https://pubmed.ncbi.nlm.nih.gov/20299481/)]

Abbreviations

ACG: Adjusted Clinical Groups score

ED: emergency department

HRA: health risk appraisal

PAM: patient activation measure

Edited by G Eysenbach; submitted 09.01.14; peer-reviewed by L Thairu, S Pit; comments to author 26.04.14; revised version received 24.06.14; accepted 12.08.14; published 02.10.14.

Please cite as:

Hibbard JH, Greene J

The Impact of an Incentive on the Use of an Online Self-Directed Wellness and Self-Management Program

J Med Internet Res 2014;16(10):e217

URL: <http://www.jmir.org/2014/10/e217/>

doi: [10.2196/jmir.3239](https://doi.org/10.2196/jmir.3239)

PMID: [25280348](https://pubmed.ncbi.nlm.nih.gov/25280348/)

©Judith H Hibbard, Jessica Greene. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 02.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Development and Preliminary Evaluation of an Internet-Based Healthy Eating Program: Randomized Controlled Trial

Katy Tapper¹, PhD; Gabriela Jiga-Boy², PhD; Gregory R Maio³, PhD; Geoffrey Haddock³, PhD; Michael Lewis⁴, PhD

¹Department of Psychology, City University London, London, United Kingdom

²Department of Psychology, Swansea University, Swansea, United Kingdom

³School of Psychology, Cardiff University, Cardiff, United Kingdom

⁴College of Engineering, Swansea University, Swansea, United Kingdom

Corresponding Author:

Katy Tapper, PhD

Department of Psychology

City University London

Social Sciences Building

Whiskin Street

London, EC1R 0JD

United Kingdom

Phone: 44 20 7040 8500

Fax: 44 20 7040 8887

Email: katy.tapper.1@city.ac.uk

Abstract

Background: The HealthValues Healthy Eating Programme is a standalone Internet-based intervention that employs a novel strategy for promoting behavior change (analyzing one's reasons for endorsing health values) alongside other psychological principles that have been shown to influence behavior. The program consists of phases targeting motivation (dietary feedback and advice, analyzing reasons for health values, thinking about health-related desires, and concerns), volition (implementation intentions with mental contrasting), and maintenance (reviewing tasks, weekly tips).

Objective: The aim was to examine the effects of the program on consumption of fruit and vegetables, saturated fat, and added sugar over a 6-month period.

Methods: A total of 82 females and 18 males were recruited using both online and print advertisements in the local community. They were allocated to an intervention or control group using a stratified block randomization protocol. The program was designed such that participants logged onto a website every week for 24 weeks and completed health-related measures. Those allocated to the intervention group also completed the intervention tasks at these sessions. Additionally, all participants attended laboratory sessions at baseline, 3 months, and 6 months. During these sessions, participants completed a food frequency questionnaire (FFQ, the Block Fat/Sugar/Fruit/Vegetable Screener, adapted for the UK), and researchers (blind to group allocation) measured their body mass index (BMI), waist-to-hip ratio (WHR), and heart rate variability (HRV).

Results: Data were analyzed using a series of ANOVA models. Per protocol analysis (n=92) showed a significant interaction for fruit and vegetable consumption ($P=.048$); the intervention group increased their intake between baseline and 6 months (3.7 to 4.1 cups) relative to the control group (3.6 to 3.4 cups). Results also showed overall reductions in saturated fat intake (20.2 to 15.6 g, $P<.001$) and added sugar intake (44.6 to 33.9 g, $P<.001$) during this period, but there were no interactions with group. Similarly, there were overall reductions in BMI (27.7 to 27.3 kg/m², $P=.001$) and WHR (0.82 to 0.81, $P=.009$), but no interactions with group. The intervention did not affect alcohol consumption, physical activity, smoking, or HRV. Data collected during the online sessions suggested that the changes in fruit and vegetable consumption were driven by the motivational and maintenance phases of the program.

Conclusions: Results suggest that the program helped individuals to increase their consumption of fruit and vegetables and to sustain this over a 6-month period. The observed reduction in fat and sugar intake suggests that monitoring behaviors over time is effective, although further research is needed to confirm this conclusion. The Web-based nature of the program makes it a potentially cost-effective way of promoting healthy eating.

KEYWORDS

social values; diet; fruit; vegetables; saturated fat; added sugar; motivation; Internet; health promotion; psychology

Introduction

A diet that is high in saturated fat and added sugars and low in fruit and vegetables is associated with a range of chronic diseases, including cardiovascular disease, cancer, and diabetes [1-5]. However, such a diet is typical for a large proportion of European and North American adults [3,6-8], and lifestyle-related diseases are now the leading cause of death globally [9]. Therefore, dietary improvement has become a priority for many Western governments [10].

One way of promoting a more healthy diet is via Internet-based intervention. This has a range of potential advantages [11], including the ability to incorporate interactive and tailored features into a program that is fully automated. This makes it a potentially cost-effective approach. Indeed, a number of fully automated Internet interventions have shown positive effects on diet. For example, compared with control groups, 4 studies have found significant reductions in fat intake up to 8 months from baseline [12-15], 3 studies have found significant increases in fruit and vegetable consumption up to 15 months [15-17], and 1 study has found a significant reduction in added sugar intake at 4 months, although not at 8 months [15].

Although these results offer a useful first step in understanding the efficacy of Internet-based health promotion interventions, most of them draw on the same set of behavior change theories to guide content development. In particular, social cognitive theory, the theory of reasoned action / planned behavior, and the transtheoretical model are frequently used [18]. Theory is a powerful tool for effective interventions [18], but these models sometimes lack empirical support and specific details about how to actually change behavior [19-21]. Additionally, they do not always encompass the latest research findings.

This paper describes the initial evaluation of a new, fully automated Internet-based healthy eating intervention: the HealthValues Healthy Eating Programme. This program differs from previous Web interventions in its use of novel behavior change techniques. In developing the HealthValues Programme, we used a more bottom-up approach, employing a selection of distinct, brief interventions that have been shown to influence behavior. There are a wide range of such techniques in the research literature, but these often fail to be translated into practice. As such, the strategies we selected can be viewed as a starting point rather than a comprehensive selection.

The first strategy involved asking individuals to spend 5 minutes thinking about why the value of health is important or unimportant to them. There is evidence that social values (eg, equality, helpfulness) often lack cognitive support. In other words, although individuals believe them to be important, they have not necessarily thought about why they are important [22]. This means that they tend to behave in accordance with the value only when it is relatively easy to do so. However, asking individuals to think about the reasons underpinning social values

can help them build cognitive support for these values and, in turn, promote more value-consistent behavior [23]. Recent research has suggested that health values also lack cognitive support, to the extent that thinking about reasons for health can have a positive influence on eating behaviors [24]. Given that this lack of cognitive support was evident across a range of social groupings and regardless of whether individuals lead healthy or unhealthy lifestyles, it suggests that this very simple strategy may be beneficial for a large number of individuals.

The second and third strategies asked individuals to spend 5 minutes considering (1) their desires and aspirations in relation to their health together with how achieving these would make them feel and (2) their concerns in relation to their health alongside how failing to avoid these would make them feel. These strategies map onto techniques commonly employed in motivational interviewing (MI) [25]. MI aligns with the principles of self-determination theory (SDT) [26] and has been shown to be effective in promoting dietary change [27]. These 2 strategies also draw on suggestions that affective messages may result in greater behavioral change than cognitive-based messages [28,29], but consistent with MI and SDT, these strategies take a nondirective approach.

The fourth strategy consists of implementation intentions with mental contrasting. Implementation intentions are specific plans of when, where, and how someone will change their behavior. They are believed to work by (1) increasing the accessibility of the situational cue that is relevant to the target behavior and (2) increasing the efficiency with which one performs the target behavior in the presence of the situational cue [30]. There is considerable evidence that implementation intentions can help promote behavior change [31,32]. In the present study, implementation intentions were employed in combination with mental contrasting. Mental contrasting involves thinking about both positive outcomes following successful behavior change as well as obstacles that might stand in the way of behavior change [33]. Mental contrasting with implementation intentions has been shown to reduce unhealthy snacking to a greater degree than either strategy in isolation [33] and has also been shown to increase fruit and vegetable consumption over a 2-year period [34].

To enhance the efficacy of the implementation intentions, we also utilized evidence about moderators by including a number of other features. These were the use of an “if...then...” format [35], use of self-formulated rather than assigned implementation intentions [36], visualization of the implementation intention [33], the formation of just 1 implementation intention at a time [37,38], emailed reminders of the implementation intention [18], the opportunity to review and modify the implementation intention in subsequent weeks [39,34], and a limited amount of tailored feedback aimed at promoting self-efficacy and autonomy [40].

The fifth strategy was the use of tailored dietary feedback in conjunction with standard health promotion advice [41,42]. Participants were provided with estimates of their intake of saturated fat, added sugar, and fruit and vegetables, along with government intake recommendations, information on the health consequences of high or low intake, and some simple strategies for adjusting one's diet. Although this component of the intervention was similar to what might be contained in an intervention with a more educational approach, an awareness of one's own diet and how it might be improved was deemed to be a prerequisite for subsequent change [43].

Finally, the program also incorporated weekly tips during the last phase. These were primarily aimed at maintaining user engagement [44] rather than promoting behavior change per se. They were designed to be light-hearted and engaging, but were also evidence-based.

Drawing on the model of action phases [45], these strategies were divided into a motivational phase (dietary feedback, reasons for health values, health-related desires and aspirations, health-related concerns) and a volitional phase (implementation intentions). This was followed by a maintenance phase during which participants could repeat or review previous tasks and information and could also access the Tip of the Week. We evaluated the program over a 6-month period through the use of laboratory-based measures taken at baseline, 3 months, and 6 months, and via weekly online measures. The intervention group was compared with a control group who completed the laboratory and online measures, but not the intervention strategies. The main aim of the study was to examine the effects of the program on different types of health-related eating behaviors, those that require engagement (eating more fruit and vegetables) and those that require disengagement (eating less saturated fat and added sugar). However, we were also interested in examining the spillover effects to other health-related behaviors (eg, physical activity, alcohol consumption, smoking) [46].

Methods

Sample Size

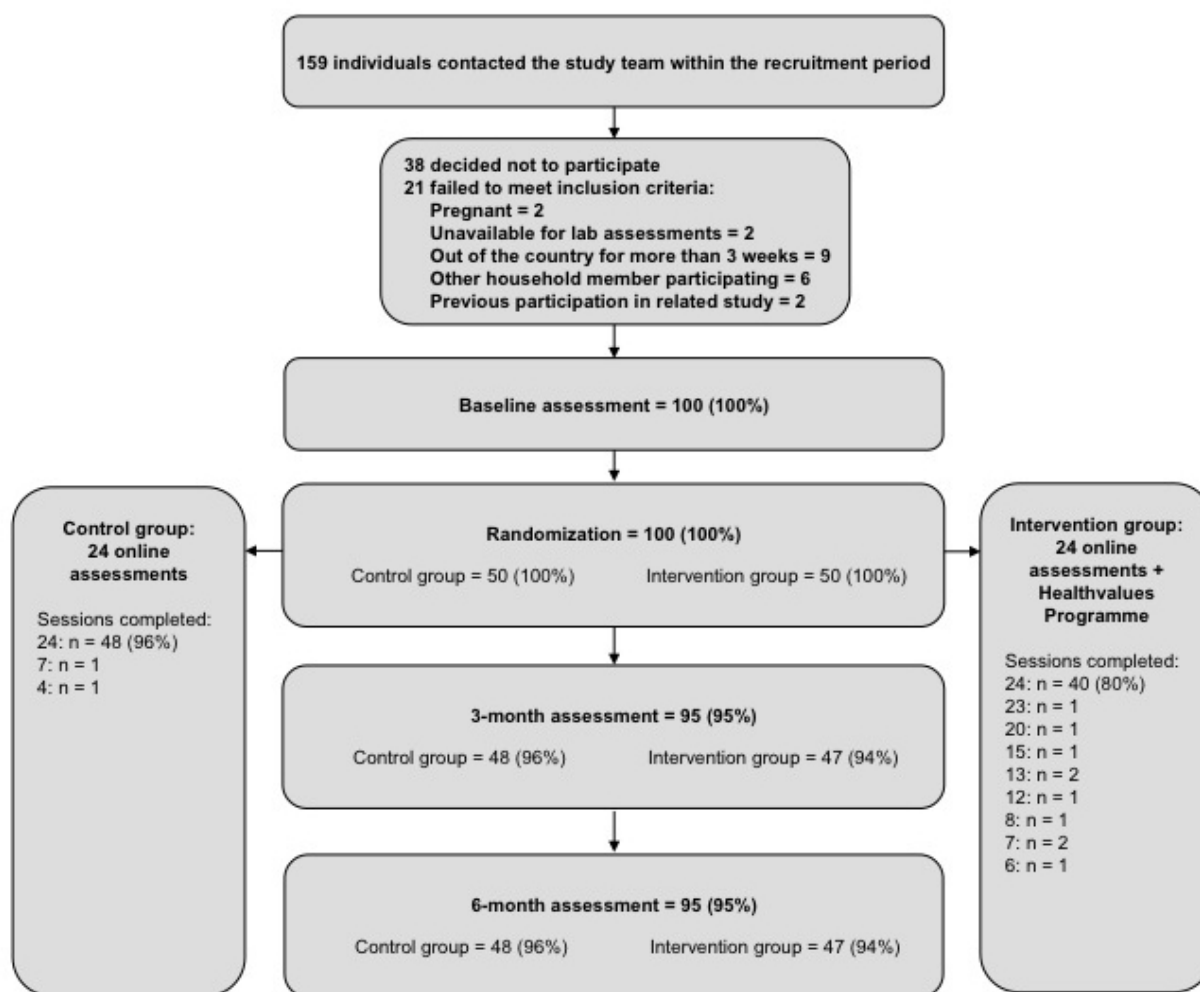
Given that this study served as an initial test of the program, there were no comparable studies on which to base sample size

calculations. That said, our sample size was informed by our previous research that examined the effects of 1 of the intervention components (thinking about reasons for values) on eating behavior over a 7-day period [24]. The eating behavior measure showed a mean difference between groups of 0.92 and a standard deviation of 1.51, meaning that at 80% power, 44 participants per group would be needed to detect a significant difference (2-tailed, $P < .05$). Assuming an attrition rate of no more than 15% [47], we concluded that a sample size of 100 would be appropriate for this trial.

Participants

Participants were recruited using both online and print advertisements in the local community. These included posters and flyers in local shops and community facilities, and advertisements on social media sites, email networks, and in local newspapers. The advertisements stated that the study team were looking for individuals to test a new online healthy eating program and noted that individuals would be reimbursed for participation. The study's website address (which included a full participant information sheet) was included in the advertisement. See [Multimedia Appendices 1 and 2](#) for study home page and information sheet.

As inclusion criteria, we stipulated that participants be aged 18 or older and able to comply with the study procedures (ie, attend the laboratory appointments and complete the weekly online sessions). Other exclusion criteria were pregnancy, being out of the country for more than 3 weeks during the study period, another household member already participating, and participation in a previous related study. A total of 159 individuals contacted the study team during the recruitment period. Of these, 38 decided not to take part or failed to respond to subsequent communications and 21 did not meet inclusion criteria. [Figure 1](#) shows the flow of participants through the study. Of the 100 participants recruited, 82 were females and 18 were males. Mean age was 39 (SD 14) years and mean body mass index (BMI) was 27.68 (SD 5.73) kg/m². A total of 23 participants were dieting to lose weight. Participants were predominantly white (93.0%, 93/100) and most had English or Welsh as a first language (94.0%, 94/100) and were well-educated (63.0%, 63/100 to degree level).

Figure 1. Flow of participants through the study.

Study Design and Procedure

The study received ethics approval from Swansea University Psychology Department Ethics Committee. Informed consent was collected by researchers at the first laboratory assessment (described subsequently). Although the study was a randomized controlled trial design, given its exploratory nature, the trial was not registered.

Laboratory measures were taken at baseline (February to April 2012), and at 3 months (May to July 2012) and 6 months (August to October 2012) postbaseline by GJB and a second research assistant, both of whom were blind to group allocation. Following baseline assessment, GJB emailed KT details of each participant's dieting status and fruit and vegetable consumption. KT then allocated participants to an intervention or control ("monitoring") group using a stratified block randomization protocol on the basis of dieting status (dieting versus nondieting) and fruit and vegetable consumption (≥ 5 a day versus < 5 a day). Block size was 2 and random numbers were generated in Excel. KT then emailed the participant details of their user ID and password and they were informed of their group allocation the first time they logged on. Although participants were not blind to group allocation, they were informed that both the "experimental" group and the "monitoring" group would

monitor eating behaviors and that this had been shown to be useful for reaching health goals. Participants in the control group were offered the opportunity to complete the program tasks at the end of the study.

All participants were asked by automated email to log onto the study website every week on 24 separate occasions to complete measures (intervention and control group) and program tasks (intervention group only). Each session could be accessed 6 days after completion of the previous session. Once the session became available, the participant was sent an email asking them to log in to complete it. Up to 3 automated reminders were emailed 2, 4, and 6 days later to participants who failed to complete the session. After completion of each session, the participant was sent an automated email thanking them and reminding them to log in again the following week. If participants failed to log in for 3 weeks, GJB attempted to contact them by phone and then email to establish whether they still wanted to participate in the online sessions and, if not, to assure them that we would still be keen for them to attend the laboratory assessments.

Each participant received £10 (approximately US \$17) for attending the first laboratory session, £25 (US \$42) for the second, and £50 (US \$84) for the third. Additionally they received £2 (US \$3) per session for completing the first 10

online sessions, £2.50 (US \$4) per session for completing the next 10 online sessions, and £5 (US \$8) per session for completing the last 4 online sessions. Thus, participants could receive up to £150 (US \$253) for completing all laboratory and online sessions. Money for completing the online sessions was given at the final laboratory assessment and amounts allocated were indicated in emails sent to prompt, remind, and thank participants. In a further effort to limit attrition, participants received small gifts (a fabric bag and a mouse pad) at the first and second laboratory assessments. These were branded with the HealthValues logo.

Measures

Outcome Measures

Primary outcome measures were intake of (1) saturated fat, (2) added sugar, and (3) fruit and vegetables. These were assessed in a laboratory using the Block Fat/Sugar/Fruit/Vegetable screener, a 55-item food frequency questionnaire (FFQ) adapted from a longer version that has been shown to have good reliability and validity [48,49]. The FFQ included questions about both frequency and quantity of intake. It was developed in North America and for our purposes adapted for use in the UK. Because the questionnaire often referred to quantities in terms of “cups,” participants were also given 4 UK measuring cups (1 cup, 1/2 cup, 1/4 cup, 1/8 cup) to assist them with their portion estimates when completing the questionnaire.

Secondary outcome measures were BMI, waist-to-hip ratio (WHR), heart rate variability (HRV), smoking status, smoking frequency, quantity of alcohol consumed, binge drinking, physical activity, dietary behaviors, and additional online assessments of saturated fat, added sugar, and fruit and vegetable intake. BMI, WHR, and HRV were assessed in the laboratory by trained researchers. These physiological measures provide an objective assessment of health status [50]. For example, HRV is a surrogate measure of cardiac control via the autonomic nervous system and can be considered to be a measure of cardiac fitness. Less favorable HRV profiles are associated with hypertension, cardiovascular disease, and aging [51], whereas physical activity has a positive effect on HRV profile [52,53]. In this study, we quantified HRV using the common statistical indexes standard deviation of cardiac (“RR”) interval (SDRR) and root mean square of successive differences (RMSSD), which reflect overall HRV and short-term (respiratory-mediated) HRV, respectively [54]. Higher scores represent better cardiac control.

Alcohol consumption was measured in the laboratory using a questionnaire designed to capture episodes of binge drinking as well as typical drinking behaviors [55]. It contained 4 items asking about frequency of consumption and number of units consumed for both usual consumption and for days when the respondent consumed larger-than-usual quantities. The questionnaire was scored by converting frequencies to drinks per week and then multiplying frequency by number of units to obtain the number of units consumed per week from usual drinking. To compute additional units consumed from larger-than-usual episodes, the usual number of units consumed was first subtracted from the larger-than-usual number of units. This gave the number of additional units consumed on these occasions. This number was then multiplied by the

larger-than-usual frequency to obtain a figure for the additional number of units consumed per week from more-than-usual drinking. The 2 figures were then added together to obtain the overall number of units consumed per week. In-line with British government recommendations, binge drinking was defined as 8 or more units per day for men and 6 or more units per day for women [56]. Where quantities consumed for either usual consumption or larger-than-usual consumption met these criteria, they were coded as an episode of binge drinking.

Smoking was assessed in the laboratory by asking participants whether they smoked cigarettes and, if yes, the number they usually smoked either per day, per week, or per month. Scores were recorded into number smoked per week.

Physical activity was assessed online at sessions 1, 8, 12, and 24 using the short version of the International Physical Activity Questionnaire (IPAQ) [57]. Participants indicated on how many days and for how long they had engaged in vigorous activity, moderate activity, and walking during the previous week. These scores were converted into total number of metabolic equivalent of task (MET) units expended per day [58].

In addition to the laboratory assessments, saturated fat, added sugar, and fruit and vegetable consumption were also assessed online at sessions 1, 8, 12, and 24 using a validated UK FFQ [59]. Respondents recorded the frequency with which they consumed 63 common food items over the previous month. The FFQ has been shown to have good test-retest reliability [60], as well as good convergent validity with 10-day weighed records [61] and with 24-hour dietary records [59]. The FFQ has also been shown to possess good construct validity [62].

To compute daily intake of saturated fat and added sugar, the proportions of these macronutrients in each of the 63 foods were calculated based on data provided by the British Food Standards Agency [63,64]. Each participant’s daily intake of each food was then computed by multiplying frequency of consumption by average portion size. Average portion sizes were based on Bingham and Day [65] and the British Food Standards Agency [64]. Finally, the quantities of saturated fat and added sugar consumed were calculated by multiplying daily intake values of the various food types by the proportion of saturated fat/added sugar in each food. These were then summed across the 63 foods to provide daily total consumption of saturated fat and added sugar for each participant.

Two additional questions were used in the calculation of fruit and vegetable consumption. These were the number of portions of fruit (excluding fruit juice), and the number of portions of vegetables (excluding potatoes, beans, and lentils) eaten on a typical day during the previous week. Examples of portions were provided. These scores were combined with scores from items relating to fruit juice and beans/lentils from the FFQ to compute daily servings of fruit and vegetables. In-line with UK guidelines, juice and beans/lentils were counted as a maximum of 1 serving a day each.

Dietary behaviors were assessed at the start of each of the 24 online sessions using a questionnaire that was developed for the project. This consisted of 17 items associated with standard dietary advice related to consumption of saturated fat, added

sugar, and fruit and vegetables (eg, reducing the number of teaspoons of sugar added to hot drinks, cereals, and desserts; replacing red meat with white meat or fish). The items were a mix of quantitative (eg, number of high fat snacks during the previous week) and categorical (eg, type of milk usually consumed). To reduce respondent burden, after the first session participants were presented with their responses from the previous session and asked to simply adjust their answers where they had made a dietary change. The questionnaire was scored by calculating the number of positive versus negative changes made since the previous session (−17 to +17).

All online questionnaires were tested for usability before the study. Questionnaires and items were presented in the same order for each participant and participants needed to complete all items before progressing to the next screen. Adaptive questioning was used for the IPAQ.

Demographic Measures

Details of participants' gender, age, level of education, and first language were collected at the first online session.

Additional Measures

Data relating to potential mediators (habits, intentions, self-efficacy, anticipated emotions), moderators (need for affect, need for cognition, behavioral approach system sensitivity, behavioral inhibition system sensitivity, environmental change), and process measures (poststudy feedback questionnaires and telephone interviews) were also collected, but these are not discussed in the present paper.

Intervention

The intervention was tested for usability before the study. At all sessions, intervention components were delivered after assessment measures. The intervention components are detailed in [Multimedia Appendix 3](#). For information purposes, [Multimedia Appendix 1](#) also shows how the components relate to Michie and colleagues' recommended taxonomy of behavior change techniques [66]. Further details of the intervention components can be obtained from the first author.

Statistical Analysis

Baseline characteristics of the 2 groups were compared using *t* tests and chi-square tests. Given the exploratory nature of the trial, intention-to-treat analyses were conducted on primary

outcomes only. Missing data were replaced by calculating the mean change from previous observations in the control group and adding or subtracting this figure from the previous observation relating to the missing data point. To examine changes in time over the 6-month period, ANOVA models with time as an independent variable were employed for the main analyses. Thus, a series of 3×2 mixed ANOVA models were used to examine the effects of the intervention on laboratory-measured intake of (1) saturated fat, (2) added sugar, and (3) fruit and vegetables. Independent variables were time (baseline, 3 months, 6 months) and group (control, intervention). There were 7 outliers (defined as greater than 3.5 SDs from the mean) and the analysis was conducted both with these unchanged and by adjusting them to 3.5 SDs from the mean.

Per protocol analysis was conducted on all primary and secondary outcomes by including only those participants who completed all 3 laboratory assessments as well as 12 or more of the 24 online sessions (for laboratory measures) or all 24 online sessions (for online measures). Although the samples for such analyses are subject to bias, they are an important means of examining intervention efficacy in exploratory trials. A series of 3 (time) × 2 (group) mixed ANOVA models were used to examine effects on laboratory-based measures whereas 4 (time) × 2 (group) ANOVA models were used for online measures. Analyses were conducted with outliers (defined as 3.5 SDs from the mean) both included and excluded. Fisher exact test was used to examine smoking status and chi-square test was used for binge drinking status.

To examine the effects of the individual intervention strategies employed in the motivational phase, change scores were calculated using the dietary behaviors questionnaire. These were computed using figures from the session in which the strategy was employed and 2 sessions later (eg, change between sessions 1 and 3, see [Multimedia Appendix 1](#) for details of strategies). Change score was then employed as the dependent variable in a 2 (condition) × 4 (strategy) mixed ANOVA.

Results

Baseline Characteristics

Analysis of baseline characteristics showed that the intervention and control groups were well matched across a range of variables (see [Table 1](#)).

Table 1. Baseline characteristics of the intervention and control groups (N=100).

Variable	Control group (n=50)	Intervention group (n=50)	P value
Gender (female), n (%)	42 (84)	41 (82)	.79 ^b
Age (years), mean (SD)	37.7 (13.2)	41.1 (14.1)	.21 ^c
BMI (kg/m ²), mean (SD)	28.1 (5.8)	27.1 (5.7)	.40 ^c
Dieting status (dieting), n (%)	11 (22)	12 (24)	.81 ^b
Education level (degree level or higher), ^a n (%)	29 (58)	34 (68)	.86 ^b
First language (English/Welsh), n (%)	49 (98)	45 (90)	.09 ^b
Ethnic background (white British), n (%)	42 (84)	34 (68)	.32 ^b

^a Highest level of educational attainment coded as GCSEs, A-levels, degree (or equivalent), still studying or other.

^b *t* test.

^c Chi-square test.

Intention-to-Treat Analyses

Descriptive and inferential statistics for intention-to-treat analyses (without outlier adjustment) are shown in Table 2. The results suggest that although both groups showed significant reductions in saturated fat and added sugar over the 6-month period, participants allocated to the intervention group did not

show greater improvements than those allocated to the control group. There was no overall change in fruit and vegetable consumption over time, but a trend toward an increase in the intervention group relative to the control group (small to medium effect size). Repeating the analyses with outlier adjustment showed near identical results.

Table 2. Means (SDs) and results from ANOVA models for intake of saturated fat, added sugar, and fruit and vegetables at baseline, 3 months, and 6 months in the intervention and control groups, for the intention-to-treat analysis.

Variable and time	Group, mean (SD)		Effects for time			Effects for time × group		
	Control (n=50)	Intervention (n=50)	<i>F</i> _{1, 98}	<i>P</i>	Partial η^2	<i>F</i> _{1, 98}	<i>P</i>	Partial η^2
Saturated fat (grams)			35.9	<.001	0.27	0.8	.36	0.01
Baseline	21.4 (8.9)	19.7 (9.6)						
3 months	17.3 (8.3)	16.1 (7.7)						
6 months	15.9 (6.6)	15.7 (9.9)						
Added sugar (grams)			8.6	.004	0.08	0.2	.62	0.00
Baseline	47.6 (34.0)	43.2 (42.0)						
3 months	36.7 (30.4)	30.3 (25.5)						
6 months	38.5 (37.6)	30.5 (37.0)						
Fruit and vegetables (cups)			0.0	.98	0.00	3.1	.08	0.03
Baseline	3.6 (1.5)	3.7 (1.7)						
3 months	3.5 (1.9)	3.8 (1.7)						
6 months	3.3 (1.5)	3.9 (1.6)						

Per Protocol Analyses

Descriptive and inferential statistics for continuous primary and secondary outcome measures collected at laboratory sessions are shown in Table 3. Over the 6-month period, participants in both groups showed comparable declines in saturated fat intake, added sugar intake, BMI, and WHR. For fruit and vegetable intake, the intervention group showed significant increases relative to the control group. Follow-up independent *t* tests indicated no difference in fruit and vegetable consumption

between the intervention and control groups at baseline and 3 months ($t_{90}=0.31, P=.78$ and $t_{90}=1.01, P=.28$, respectively), but significantly greater intake in the intervention group at 6 months ($t_{90}=2.30, P=.02$). For the RMSSD HRV measure there was a trend toward a significant group × time interaction, but no main effect of time. SDRR HRV and total alcohol intake did not change over time and were not influenced by group status. The same pattern of results occurred when these analyses were repeated but with outliers excluded.

Table 3. Means (SDs) and results from ANOVA models for laboratory-assessed primary and secondary outcomes at baseline, 3 months, and 6 months in the intervention and control groups, for the per protocol analyses.

Variable and time	Group, mean (SD)		Effects for time		Effects for time × group			
	Control (n=47) ^a	Intervention (n=45) ^b	<i>F</i> _{1, 90} ^c	<i>P</i>	Partial η^2	<i>F</i> _{1, 90} ^c	<i>P</i>	Partial η^2
Saturated fat (grams)			28.7	<.001	0.24	1.2	.23	0.01
Baseline	21.0 (8.9)	19.3 (8.9)						
3 months	16.7 (8.0)	16.2 (7.3)						
6 months	15.5 (6.4)	15.7 (9.6)						
Added sugar (grams)			7.2	.009	0.07	0.1	.76	0.00
Baseline	46.7 (34.3)	42.3 (43.0)						
3 months	35.8 (30.5)	30.4 (26.1)						
6 months	37.2 (38.1)	30.4 (38.5)						
Fruit and vegetables (cups)			0.3	.57	0.00	4.0	.048	0.04
Baseline	3.6 (1.4)	3.7 (1.7)						
3 months	3.4 (1.7)	3.8 (1.7)						
6 months	3.4 (1.5)	4.1 (1.6)						
Alcohol (units per week)			1.6	.20	0.02	0.2	.69	0.00
Baseline	6.4 (5.6)	6.3 (6.2)						
3 months	6.8 (7.2)	6.7 (6.9)						
6 months	7.2 (7.5)	6.7 (7.3)						
BMI (kg/m²)			11.2	.001	0.11	0.1	.93	0.00
Baseline	28.4 (5.8)	27.0 (5.9)						
3 months	28.3 (5.9)	26.8 (5.7)						
6 months	28.0 (5.9)	26.6 (5.9)						
WHR			7.2	.009	0.07	0.0	.71	0.00
Baseline	0.82 (0.09)	0.82 (0.09)						
3 months	0.81 (0.09)	0.82 (0.09)						
6 months	0.81 (0.08)	0.81 (0.08)						
HRV: SDRR (ms)			1.4	.25	0.02	2.0	.13	0.02
Baseline	45.0 (20.1)	49.6 (19.7)						
3 months	46.4 (20.1)	47.8 (18.7)						
6 months	46.1 (17.9)	43.1 (15.2)						
HRV: RMSSD (ms)			1.4	.24	0.02	2.9	.06	0.03
Baseline	28.9 (14.6)	33.1 (19.6)						
3 months	19.3 (15.3)	30.5 (16.3)						
6 months	30.2 (15.4)	25.8 (12.9)						

^a For alcohol consumption, n=46 due to questionnaire completion error.

^b For alcohol consumption, n=44 due to questionnaire completion error.

^c For alcohol consumption, *F*_{1, 88}.

For smoking status, there were 91 participants who provided data on smoking at all 3 laboratory assessments and completed at least 12 of the online sessions. At each of the 3 time points there was no difference in the proportion of smokers in the experimental group compared to the control group at baseline

(control: n=6, experimental: n=2, *P*=.27), 3 months (control: n=4, experimental: n=1, *P*=.36), and at 6 months (control: n=4, experimental: n=3, *P*>.99). Smoking frequency was not analyzed due to the small number of smokers in the sample.

Analysis of binge drinking included 90 participants who provided data on alcohol consumption at all 3 laboratory assessments and completed at least 12 of the online sessions. Again, at each of the 3 time points, there was no difference in the proportion of individuals who engaged in binge drinking in the experimental group compared to the control group at baseline (control: n=25, experimental: n=23; $\chi^2_1=0.0, P=.84$), 3 months (control: n=23, experimental: n=17; $\chi^2_1=1.1, P=.28$) and 6 months (control: n=20, experimental: n=17; $\chi^2_1=0.2, P=.64$).

Descriptive and inferential statistics for secondary outcome measures collected during the online sessions are shown in Table 4. Consistent with laboratory assessments, these show there were significant reductions in intake of saturated fat and added sugar over time, but that the extent of these reductions

did not differ between intervention and control groups. Also consistent with laboratory assessments, the results show an increase in fruit and vegetable consumption among the intervention group relative to the control group. This was coupled with an overall increase in fruit and vegetable consumption over time. Follow-up independent *t* tests indicated no difference in fruit and vegetable consumption between the intervention and control groups at sessions 1, 8, and 12 ($t_{86}=0.19, P=.85$; $t_{86}=1.64, P=.11$; $t_{86}=1.48, P=.14$, respectively), but significantly greater intake in the intervention group at session 24 ($t_{86}=2.45, P=.02$). Additionally, the results showed no significant change in physical activity over time and no effect of the intervention on physical activity. The same pattern of results occurred when these analyses were repeated with outliers excluded.

Table 4. Means (SDs) and results from ANOVA models for secondary outcomes assessed online at sessions 1, 8, 12, and 24 in the intervention and control groups, for the per protocol analyses.

Variable and session	Group, mean (SD)		Effects for time			Effects for time × group		
	Control (n=48) ^a	Intervention (n=40) ^b	<i>F</i> _{1, 86} ^c	<i>P</i>	Partial η^2	<i>F</i> _{1, 86} ^c	<i>P</i>	Partial η^2
Saturated fat (grams)			7.8	.006	0.08	0.6	.43	0.01
1	24.4 (9.9)	26.0 (15.4)						
8	22.3 (10.6)	21.4 (13.0)						
12	21.2 (10.4)	21.7 (11.4)						
24	22.4 (10.0)	21.5 (9.1)						
Added sugar (grams)			8.41	.005	0.10	2.0	.16	0.02
1	47.8 (43.6)	57.32 (74.5)						
8	34.4 (22.7)	34.4 (32.3)						
12	31.7 (21.4)	32.1 (25.3)						
24	39.8 (27.0)	31.8 (19.4)						
Fruit and vegetables (portions)			5.6	.02	0.06	5.5	.02	0.06
1	4.9 (2.1)	5.0 (2.0)						
8	5.2 (2.4)	6.0 (2.3)						
12	5.3 (2.8)	6.1 (2.2)						
24	4.9 (2.3)	6.2 (2.7)						
Physical activity (METS per week)			0.2	.67	0.00	0.2	.69	0.00
1	2857 (2320)	2432 (1626)						
8	2534 (2290)	2138 (1522)						
12	2932 (4270)	2420 (1966)						
24	2985 (3525)	2350 (2344)						

^a For physical activity n=39 due to participants coding “don’t know.”

^b For physical activity n=37 due to participants coding “don’t know.”

^c For physical activity, *F*_{1, 74}.

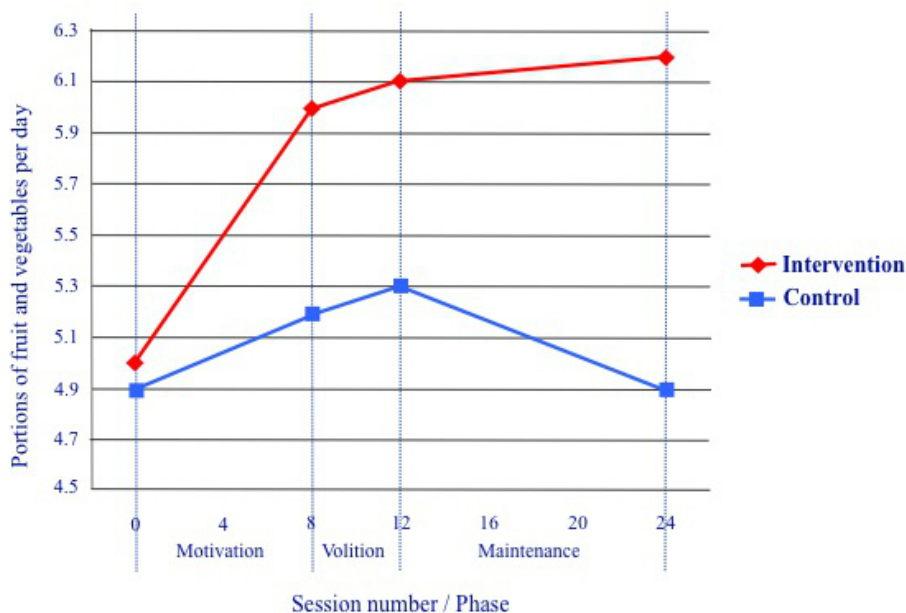
Figure 2 shows levels of fruit and vegetable consumption in the intervention and control groups at the start and end of each of the 3 program phases. As noted previously, follow-up analyses indicated that significant differences between intervention and control groups occurred at the fourth measurement point only

(ie, session 24, the end of the third phase, $t_{86}=2.45, P=.02$). These results, together with Figure 2, suggest that the most likely explanation for this effect is that it was driven primarily by the combination of motivation and maintenance phases. However, it is also possible that the maintenance phase played

no part in the changes, but that the differences at session 24 were a result of the motivational phase continuing to exert effects over the 6-month period. Additionally, the data suggest that (in its position within the intervention) the volitional phase had no immediate impact (although a delayed impact cannot be ruled out). The pattern of results from the per protocol analysis

were unchanged after repeating the analysis with only the intervention participants who had formed at least 1 volitional phase implementation intention related to the relevant outcome measure (fruit and vegetables: $n=24$; saturated fat: $n=30$; added sugar: $n=32$).

Figure 2. Portions of fruit and vegetables consumed in the intervention and control groups at the start and end of each program phase.



Effects of Individual Strategies Employed in the Motivational Phase

For analysis of motivational phase strategies, all participants who completed the first 9 online sessions were included (control: $n=47$; intervention: $n=46$). Because fruit and vegetable consumption was improved by the intervention, we conducted exploratory analyses examining changes in fruit and vegetable consumption in the intervention and control groups in the 2-week period following the delivery of each of the 4 different program components (see Figure 3). There was no main effect of strategy ($F_{1, 91}=0.53$, $P=.47$, partial eta-squared=0.01) or

condition ($F_{1, 91}=0.87$, $P=.47$, partial eta-squared=0.01) and no significant interaction between strategy and condition ($F_{1, 91}=2.88$, $P=.09$, partial eta-squared=0.03), although the latter results are marginal. These results suggest that the increases in fruit and vegetable consumption seen in the intervention group were brought about by a combination of intervention components in both the motivational and maintenance phases. Figure 3 suggests that the strategy employed in session 1 (tailored feedback and advice) may have been particularly useful in eliciting change, although further research is needed to confirm this.

Figure 3. Portions of fruit and vegetables consumed in the intervention and control groups during the motivational phase.

Discussion

Results of the per protocol analysis indicated that the HealthValues Healthy Eating Programme brought about significant increases in fruit and vegetable consumption relative to a control group. These equated to approximately 0.75 cups, or 1.3 portions of the recommended 5 or more portions per day. The results also suggested that these increases were brought about primarily by strategies employed in the motivational and maintenance phases of the program, rather than the implementation intentions employed in the volitional phase. Thus, it may be that low fruit and vegetable consumption among this particular group was limited primarily by motivation rather than any difficulties in implementing the behavior; when we increased motivation, it had a direct effect on consumption.

In contrast, although the program was associated with a decrease in saturated fat and added sugar consumption, these effects were comparable to those found in the control condition. Unlike increasing fruit and vegetable intake, which involves introducing additional foods into the diet, reducing fat and sugar entails cutting back. As such, intake may be influenced by additional factors that may not be as amenable to motivational strategies. In particular, consumption of high fat and sugar foods may be habitual and carried out with a degree of automaticity [67,68]. Because habits tend to be resistant to changes in attitude [69], motivational strategies alone may be ineffective in eliciting a reduction in these forms of consumptive behavior. Additionally, foods that are high in fat and sugar may be the target of cravings [70]. Again, motivational strategies may not be sufficient to overcome such cravings. Thus, techniques specifically designed to target habits and cravings might usefully be incorporated into future versions of the program.

The results did, however, show overall reductions in intake of saturated fat and added sugar among both groups by approximately 4.7 and 11.4 grams per day, respectively. These findings are consistent with the physiological data that showed

significant reductions in BMI and WHR. Given that our recruitment method targeted individuals who wanted to improve their diet, it is possible that these changes would have occurred even in the absence of study participation. However, this seems unlikely given the general trend for weight to increase over time [71] and the fact that these data were collected over an extended (6-month) period. Instead, we would suggest that these changes might have been brought about by the monitoring component of the study, particularly the weekly brief diet questionnaire that mapped directly onto dietary advice. This questionnaire may have increased participants' knowledge of how to cut back on fat and sugar. It may also have increased attitude accessibility, the ease with which attitudes are retrieved from memory [72]. If intake of fat and sugar are determined by relatively weak habits, increased accessibility of negative attitudes toward fat and sugar may have been sufficient to disrupt automatic behaviors. Further research is needed to confirm this. It would also be important to control for the effects of researcher contact. In the current study, it is possible that the laboratory assessments, together with the incentives, may have inadvertently led to participants trying to please the researchers. These may have contributed in some small part to the overall reductions in fat and sugar intake.

The absence of effects for implementation intentions are at odds with previous non-Internet interventions [34,73], but are in-line with several other Internet-based studies [74-77]. One explanation is that participants had already formed action plans in response to the monitoring component of the study, making it difficult for the implementation intentions to bring about further change. This interpretation is consistent with other research showing implementation intentions to be less effective among individuals who are already good at action planning [78]. It also has implications for the development of interventions; because longer interventions may increase rates of drop out, it is important that all strategies employed make a unique contribution to behavior change. However, an alternative explanation is that the fruit and vegetable-related implementation

intentions helped sustain behavior change [34]. A weakness of the current study is that it is unable to distinguish between these possibilities or to identify with precision the components that are responsible for the effects. In future work, it would be helpful to compare different versions of the program to help determine which components are important and which may be redundant.

The benefits of participation did not generalize to behaviors that were not directly targeted by the program; there were no significant spillover effects on levels of physical activity, alcohol consumption, smoking, or HRV, either between groups or over time. Although some research has suggested that health improvements may show spillover effects to other health-related behaviors [46], the results of this study suggest that effects are restricted to behaviors that are targeted.

In future research, it would be important to trial the program in the absence of incentives for session completion. Given the high rates of attrition in online interventions [79], we incorporated these incentives to enable a proper initial evaluation of the program. However, a trial without these incentives would help

indicate natural attrition and allow for calculations of cost-effectiveness.

It is also important to examine the effects of the program with different populations. In the current study, we recruited participants who were interested in improving their diet. Thus, they were a group who were already reasonably motivated (as indicated by a baseline mean of 4.16 on a scale of 1 to 5 for intention to eat a healthy diet). It is possible that the motivational strategies would have been more effective among a less motivated group of individuals who might be accessed via workplace settings, for example.

In conclusion, the HealthValues Healthy Eating Programme significantly increased fruit and vegetable consumption among users. Future research comparing different versions of the program should help to identify more accurately the elements that were responsible for this effect. It seems likely that the monitoring component of the study also brought about reductions in intake of saturated fat and added sugar, although further research is needed to confirm this. Given that the program is fully automated, it represents a potentially cost-effective way of promoting healthy eating.

Acknowledgments

The research was funded by the Economic and Social Research Council. We thank Clare Clement for help with data collection, Paul Tapper for technical support, and Jennie Davies for nutritional advice.

Conflicts of Interest

None declared.

Multimedia Appendix 1

HealthValues Healthy Eating Programme homepage.

[[JPG File, 502KB - jmir_v16i10e231_app1.jpg](#)]

Multimedia Appendix 2

Study information sheet provided to participants.

[[PDF File \(Adobe PDF File\), 62KB - jmir_v16i10e231_app2.pdf](#)]

Multimedia Appendix 3

Intervention components used in the HealthValues Programme together with most closely aligned categories of behaviour change technique according to the Behavior Change Technique Taxonomy.[66] Shaded rows indicate optional components.

[[PDF File \(Adobe PDF File\), 11KB - jmir_v16i10e231_app3.pdf](#)]

Multimedia Appendix 4

CONSORT-EHEALTH checklist V1.6.2 [80].

[[PDF File \(Adobe PDF File\), 987KB - jmir_v16i10e231_app4.pdf](#)]

References

1. Esposito K, Kastorini CM, Panagiotakos DB, Giugliano D. Prevention of type 2 diabetes by dietary patterns: a systematic review of prospective studies and meta-analysis. *Metab Syndr Relat Disord* 2010 Dec;8(6):471-476. [doi: [10.1089/met.2010.0009](https://doi.org/10.1089/met.2010.0009)] [Medline: [20958207](https://pubmed.ncbi.nlm.nih.gov/20958207/)]

2. Hooper L, Summerbell CD, Thompson R, Sills D, Roberts FG, Moore HJ, et al. Reduced or modified dietary fat for preventing cardiovascular disease. *Cochrane Database Syst Rev* 2012;5:CD002137. [doi: [10.1002/14651858.CD002137.pub3](https://doi.org/10.1002/14651858.CD002137.pub3)] [Medline: [22592684](https://pubmed.ncbi.nlm.nih.gov/22592684/)]
3. Scarborough P, Nnoaham KE, Clarke D, Capewell S, Rayner M. Modelling the impact of a healthy diet on cardiovascular disease and cancer mortality. *J Epidemiol Community Health* 2012 May;66(5):420-426. [doi: [10.1136/jech.2010.114520](https://doi.org/10.1136/jech.2010.114520)] [Medline: [21172796](https://pubmed.ncbi.nlm.nih.gov/21172796/)]
4. Greenwood DC, Threapleton DE, Evans CE, Cleghorn CL, Nykjaer C, Woodhead C, et al. Glycemic index, glycemic load, carbohydrates, and type 2 diabetes: systematic review and dose-response meta-analysis of prospective studies. *Diabetes Care* 2013 Dec;36(12):4166-4171. [doi: [10.2337/dc13-0325](https://doi.org/10.2337/dc13-0325)] [Medline: [24265366](https://pubmed.ncbi.nlm.nih.gov/24265366/)]
5. Oyebode O, Gordon-Dseagu V, Walker A, Mindell JS. Fruit and vegetable consumption and all-cause, cancer and CVD mortality: analysis of Health Survey for England data. *J Epidemiol Community Health* 2014 Sep;68(9):856-862 [FREE Full text] [doi: [10.1136/jech-2013-203500](https://doi.org/10.1136/jech-2013-203500)] [Medline: [24687909](https://pubmed.ncbi.nlm.nih.gov/24687909/)]
6. Schmidhuber J, Traill WB. The changing structure of diets in the European Union in relation to healthy eating guidelines. *PHN* 2007 Jan 2;9(05):584-595. [doi: [10.1079/PHN2005844](https://doi.org/10.1079/PHN2005844)]
7. Casagrande SS, Wang Y, Anderson C, Gary TL. Have Americans increased their fruit and vegetable intake? The trends between 1988 and 2002. *Am J Prev Med* 2007 Apr;32(4):257-263. [doi: [10.1016/j.amepre.2006.12.002](https://doi.org/10.1016/j.amepre.2006.12.002)] [Medline: [17383556](https://pubmed.ncbi.nlm.nih.gov/17383556/)]
8. McCabe-Sellers BJ, Bowman S, Stuff JE, Champagne CM, Simpson PM, Bogle ML. Assessment of the diet quality of US adults in the Lower Mississippi Delta. *Am J Clin Nutr* 2007 Sep;86(3):697-706 [FREE Full text] [Medline: [17823435](https://pubmed.ncbi.nlm.nih.gov/17823435/)]
9. World Health Organization. *Global Status Report on Noncommunicable Diseases 2010*. Geneva: World Health Organization; 2011.
10. Department of Health. *Healthy Lives, Healthy People: A Call to Action on Obesity in England*. London: Crown Copyright; 2011.
11. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J Med Internet Res* 2013 Jul;15(7):e146 [FREE Full text] [doi: [10.2196/jmir.2665](https://doi.org/10.2196/jmir.2665)] [Medline: [23859884](https://pubmed.ncbi.nlm.nih.gov/23859884/)]
12. De Bourdeaudhuij I, Stevens V, Vandelandotte C, Brug J. Evaluation of an interactive computer-tailored nutrition intervention in a real-life setting. *Ann Behav Med* 2007 Feb;33(1):39-48. [doi: [10.1207/s15324796abm3301_5](https://doi.org/10.1207/s15324796abm3301_5)] [Medline: [17291169](https://pubmed.ncbi.nlm.nih.gov/17291169/)]
13. Kroeze W, Oenema A, Campbell M, Brug J. The efficacy of Web-based and print-delivered computer-tailored interventions to reduce fat intake: results of a randomized, controlled trial. *J Nutr Educ Behav* 2008 Jul;40(4):226-236. [doi: [10.1016/j.jneb.2007.09.008](https://doi.org/10.1016/j.jneb.2007.09.008)] [Medline: [18565463](https://pubmed.ncbi.nlm.nih.gov/18565463/)]
14. Oenema A, Brug J, Dijkstra A, de Weerd I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. *Ann Behav Med* 2008 Apr;35(2):125-135. [doi: [10.1007/s12160-008-9023-1](https://doi.org/10.1007/s12160-008-9023-1)] [Medline: [18363076](https://pubmed.ncbi.nlm.nih.gov/18363076/)]
15. Sternfeld B, Block C, Quesenberry CP, Block TJ, Husson G, Norris JC, et al. Improving diet and physical activity with ALIVE: a worksite randomized trial. *Am J Prev Med* 2009 Jun;36(6):475-483. [doi: [10.1016/j.amepre.2009.01.036](https://doi.org/10.1016/j.amepre.2009.01.036)] [Medline: [19460655](https://pubmed.ncbi.nlm.nih.gov/19460655/)]
16. 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](https://doi.org/10.2105/AJPH.2008.154468)] [Medline: [20019315](https://pubmed.ncbi.nlm.nih.gov/20019315/)]
17. Greene GW, White AA, Hoerr SL, Lohse B, Schembre SM, Riebe D, et al. Impact of an online healthful eating and physical activity program for college students. *Am J Health Promot* 2012;27(2):e47-e58. [doi: [10.4278/ajhp.110606-QUAN-239](https://doi.org/10.4278/ajhp.110606-QUAN-239)] [Medline: [23113786](https://pubmed.ncbi.nlm.nih.gov/23113786/)]
18. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
19. Michie S, Abraham C. Interventions to change health behaviours: evidence-based or evidence-inspired? *Psychology & Health* 2004 Feb;19(1):29-49. [doi: [10.1080/0887044031000141199](https://doi.org/10.1080/0887044031000141199)]
20. Sniehotta FF. Towards a theory of intentional behaviour change: plans, planning, and self-regulation. *Br J Health Psychol* 2009 May;14(Pt 2):261-273. [doi: [10.1348/135910708X389042](https://doi.org/10.1348/135910708X389042)] [Medline: [19102817](https://pubmed.ncbi.nlm.nih.gov/19102817/)]
21. Sniehotta FF, Pesseau J, Araújo-Soares V. Time to retire the theory of planned behaviour. *Health Psychol Rev* 2014 Jan;8(1):1-7. [doi: [10.1080/17437199.2013.869710](https://doi.org/10.1080/17437199.2013.869710)] [Medline: [25053004](https://pubmed.ncbi.nlm.nih.gov/25053004/)]
22. Maio GR, Olson JM. Values as truisms: evidence and implications. *J Per Soc Psychol* 1998;74:294-311.
23. Maio GR, Olson JM, Allen L, Bernard MM. Addressing discrepancies between values and behavior: the motivating effect of reasons. *Journal of Experimental Social Psychology* 2001 Mar;37(2):104-117. [doi: [10.1006/jesp.2000.1436](https://doi.org/10.1006/jesp.2000.1436)]
24. Tapper K, Jiga-Boy GM, Haddock G, Maio GR, Valle C. Motivating health behaviour change: provision of cognitive support for health values. *The Lancet* 2012 Nov;380:S71 [FREE Full text] [doi: [10.1016/S0140-6736\(13\)60427-4](https://doi.org/10.1016/S0140-6736(13)60427-4)]
25. Rollnick S, Miller WR. What is Motivational Interviewing? *Behav Cognit Psychother* 2009 Jun 16;23(04):325. [doi: [10.1017/S135246580001643X](https://doi.org/10.1017/S135246580001643X)]
26. 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](https://pubmed.ncbi.nlm.nih.gov/11392867/)]

27. Martins RK, McNeil DW. Review of Motivational Interviewing in promoting health behaviors. *Clin Psychol Rev* 2009 Jun;29(4):283-293. [doi: [10.1016/j.cpr.2009.02.001](https://doi.org/10.1016/j.cpr.2009.02.001)] [Medline: [19328605](https://pubmed.ncbi.nlm.nih.gov/19328605/)]
28. Conner M, Godin G, Norman P, Sheeran P. Using the question-behavior effect to promote disease prevention behaviors: two randomized controlled trials. *Health Psychol* 2011 May;30(3):300-309. [doi: [10.1037/a0023036](https://doi.org/10.1037/a0023036)] [Medline: [21553974](https://pubmed.ncbi.nlm.nih.gov/21553974/)]
29. Lawton R, Conner M, McEachan R. Desire or reason: predicting health behaviors from affective and cognitive attitudes. *Health Psychol* 2009 Jan;28(1):56-65. [doi: [10.1037/a0013424](https://doi.org/10.1037/a0013424)] [Medline: [19210018](https://pubmed.ncbi.nlm.nih.gov/19210018/)]
30. Webb TL, Sheeran P. How do implementation intentions promote goal attainment? A test of component processes. *Journal of Experimental Social Psychology* 2007 Mar;43(2):295-302. [doi: [10.1016/j.jesp.2006.02.001](https://doi.org/10.1016/j.jesp.2006.02.001)]
31. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology* 2006;38:69-111. [doi: [10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)]
32. Adriaanse MA, Vinkers CD, De Ridder DT, Hox JJ, De Wit JB. Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. *Appetite* 2011 Feb;56(1):183-193 [FREE Full text] [doi: [10.1016/j.appet.2010.10.012](https://doi.org/10.1016/j.appet.2010.10.012)] [Medline: [21056605](https://pubmed.ncbi.nlm.nih.gov/21056605/)]
33. Adriaanse MA, Oettingen G, Gollwitzer PM, Hennes EP, de Ridder DTD, de Wit JBF. When planning is not enough: Fighting unhealthy snacking habits by mental contrasting with implementation intentions (MCII). *Eur J Soc Psychol* 2010 Nov 24;40(7):1277-1293. [doi: [10.1002/ejsp.730](https://doi.org/10.1002/ejsp.730)]
34. Stadler G, Oettingen G, Gollwitzer PM. Intervention effects of information and self-regulation on eating fruits and vegetables over two years. *Health Psychol* 2010 May;29(3):274-283. [doi: [10.1037/a0018644](https://doi.org/10.1037/a0018644)] [Medline: [20496981](https://pubmed.ncbi.nlm.nih.gov/20496981/)]
35. Oettingen G, Honig G, Gollwitzer PM. Effective self-regulation of goal attainment. *International Journal of Educational Research* 2000;33:705-732.
36. Adriaanse MA, de Ridder DTD, de Wit JBF. Finding the critical cue - implementation intentions to change one's diet work best when tailored to personally relevant reasons for unhealthy eating. *Pers Soc Psychol Bull* 2009;35:60-71. [doi: [10.1016/j.appet.2010.10.012](https://doi.org/10.1016/j.appet.2010.10.012)]
37. Dalton AN, Spiller SA. Too much of a good thing: the benefits of implementation intentions depend on the number of goals. *Journal of Consumer Research* 2012 Oct;39(3):600-614. [doi: [10.1086/664500](https://doi.org/10.1086/664500)]
38. Verhoeven AAC, Adriaanse MA, de Ridder DTD, de Vet E, Fennis BM. Less is more: the effect of multiple implementation intentions targeting unhealthy snacking habits. *Eur J Soc Psychol* 2013 Jul 02;43(5):344-354. [doi: [10.1002/ejsp.1963](https://doi.org/10.1002/ejsp.1963)]
39. Chapman J, Armitage CJ. Evidence that boosters augment the long-term impact of implementation intentions on fruit and vegetable intake. *Psychol Health* 2010 Mar;25(3):365-381. [doi: [10.1080/08870440802642148](https://doi.org/10.1080/08870440802642148)] [Medline: [20204966](https://pubmed.ncbi.nlm.nih.gov/20204966/)]
40. Koestner R, Horberg EJ, Gaudreau P, Powers T, Di Dio P, Bryan C, et al. Bolstering implementation plans for the long haul: the benefits of simultaneously boosting self-concordance or self-efficacy. *Pers Soc Psychol Bull* 2006 Nov;32(11):1547-1558. [doi: [10.1177/0146167206291782](https://doi.org/10.1177/0146167206291782)] [Medline: [17030894](https://pubmed.ncbi.nlm.nih.gov/17030894/)]
41. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. *Ann Behav Med* 2006 Jun;31(3):205-223. [doi: [10.1207/s15324796abm3103_2](https://doi.org/10.1207/s15324796abm3103_2)] [Medline: [16700634](https://pubmed.ncbi.nlm.nih.gov/16700634/)]
42. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010 Sep;51(3-4):214-221 [FREE Full text] [doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004)] [Medline: [20558196](https://pubmed.ncbi.nlm.nih.gov/20558196/)]
43. Weinstein ND. The precaution adoption process. *Health Psychol* 1988;7(4):355-386. [Medline: [3049068](https://pubmed.ncbi.nlm.nih.gov/3049068/)]
44. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *J Med Internet Res* 2011 Jan;13(1):e2 [FREE Full text] [doi: [10.2196/jmir.1639](https://doi.org/10.2196/jmir.1639)] [Medline: [21212045](https://pubmed.ncbi.nlm.nih.gov/21212045/)]
45. Heckhausen H. *Motivation and Action*. Berlin: Springer-Verlag; 1991.
46. Mata J, Silva MN, Vieira PN, Carraça EV, Andrade AM, Coutinho SR, et al. Motivational "spill-over" during weight control: increased self-determination and exercise intrinsic motivation predict eating self-regulation. *Health Psychol* 2009 Nov;28(6):709-716. [doi: [10.1037/a0016764](https://doi.org/10.1037/a0016764)] [Medline: [19916639](https://pubmed.ncbi.nlm.nih.gov/19916639/)]
47. Tapper K, Shaw C, Ilsley J, Hill AJ, Bond FW, Moore L. Exploratory randomised controlled trial of a mindfulness-based weight loss intervention for women. *Appetite* 2009 Apr;52(2):396-404. [doi: [10.1016/j.appet.2008.11.012](https://doi.org/10.1016/j.appet.2008.11.012)] [Medline: [19101598](https://pubmed.ncbi.nlm.nih.gov/19101598/)]
48. Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol* 1990;43(12):1327-1335. [Medline: [2254769](https://pubmed.ncbi.nlm.nih.gov/2254769/)]
49. Mares-Perlman JA, Klein BE, Klein R, Ritter LL, Fisher MR, Freudenheim JL. A diet history questionnaire ranks nutrient intakes in middle-aged and older men and women similarly to multiple food records. *J Nutr* 1993 Mar;123(3):489-501 [FREE Full text] [Medline: [8463852](https://pubmed.ncbi.nlm.nih.gov/8463852/)]
50. Lewis MJ. Heart rate variability analysis: a tool to assess cardiac autonomic function. *Comput Inform Nurs* 2005;23(6):335-341. [Medline: [16292049](https://pubmed.ncbi.nlm.nih.gov/16292049/)]
51. Gang Y, Malik M. Heart rate variability in critical care medicine. *Curr Opin Crit Care* 2002 Oct;8(5):371-375. [Medline: [12357103](https://pubmed.ncbi.nlm.nih.gov/12357103/)]
52. Melanson EL. Resting heart rate variability in men varying in habitual physical activity. *Med Sci Sports Exerc* 2000 Nov;32(11):1894-1901. [Medline: [11079519](https://pubmed.ncbi.nlm.nih.gov/11079519/)]

53. Lewis MJ, Short AL. Exercise and cardiac regulation: what can electrocardiographic time series tell us? *Scand J Med Sci Sports* 2010 Dec;20(6):794-804. [doi: [10.1111/j.1600-0838.2010.01150.x](https://doi.org/10.1111/j.1600-0838.2010.01150.x)] [Medline: [20561274](https://pubmed.ncbi.nlm.nih.gov/20561274/)]
54. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation* 1996 Mar 1;93(5):1043-1065 [FREE Full text] [Medline: [8598068](https://pubmed.ncbi.nlm.nih.gov/8598068/)]
55. Cox W. Alcohol use questionnaire. Bangor, UK: Bangor University, Wales, UK; 2003.
56. Drink Aware. URL: <http://www.drinkaware.co.uk/understand-your-drinking/is-your-drinking-a-problem/binge-drinking> [accessed 2014-05-08] [WebCite Cache ID 6PPrYIPWu]
57. Hallal PC, Victora CG. Reliability and validity of the International Physical Activity Questionnaire (IPAQ). *Med Sci Sports Exerc* 2004 Mar;36(3):556. [Medline: [15076800](https://pubmed.ncbi.nlm.nih.gov/15076800/)]
58. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003 Aug;35(8):1381-1395. [doi: [10.1249/01.MSS.0000078924.61453.FB](https://doi.org/10.1249/01.MSS.0000078924.61453.FB)] [Medline: [12900694](https://pubmed.ncbi.nlm.nih.gov/12900694/)]
59. Margetts BM, Cade JE, Osmond C. Comparison of a food frequency questionnaire with a diet record. *Int J Epidemiol* 1989 Dec;18(4):868-873. [Medline: [2621023](https://pubmed.ncbi.nlm.nih.gov/2621023/)]
60. Armitage CJ, Conner M. The theory of planned behaviour: assessment of predictive validity and 'perceived control'. *British Journal of Social Psychology* 1999;38:35-54. [doi: [10.1348/014466699164022](https://doi.org/10.1348/014466699164022)]
61. Thompson RL, Margetts BM. Comparison of a food frequency questionnaire with a 10-day weighed record in cigarette smokers. *Int J Epidemiol* 1993 Oct;22(5):824-833. [Medline: [8282461](https://pubmed.ncbi.nlm.nih.gov/8282461/)]
62. Armitage CJ, Conner M. Efficacy of a minimal intervention to reduce fat intake. *Soc Sci Med* 2001 May;52(10):1517-1524. [Medline: [11314848](https://pubmed.ncbi.nlm.nih.gov/11314848/)]
63. Foods Standards Agency. McCance and Widdowson's the composition of foods, sixth summary edition. Cambridge: Royal Society of Chemistry; 2002.
64. Food Standards Agency. Manual of Nutrition. London: Stationery Office; 2008.
65. Bingham S, Day K. Average portion weights of foods consumed by a randomly selected British population sample. *Hum Nutr Appl Nutr* 1987 Aug;41(4):258-264. [Medline: [3667343](https://pubmed.ncbi.nlm.nih.gov/3667343/)]
66. 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;46(1):81-95. [doi: [10.1007/s12160-013-9486-6](https://doi.org/10.1007/s12160-013-9486-6)]
67. van't Riet J, Sijtsma SJ, Dagevos H, De Bruijn GJ. The importance of habits in eating behaviour. An overview and recommendations for future research. *Appetite* 2011 Dec;57(3):585-596. [doi: [10.1016/j.appet.2011.07.010](https://doi.org/10.1016/j.appet.2011.07.010)] [Medline: [21816186](https://pubmed.ncbi.nlm.nih.gov/21816186/)]
68. Verhoeven AA, Adriaanse MA, Evers C, de Ridder DT. The power of habits: unhealthy snacking behaviour is primarily predicted by habit strength. *Br J Health Psychol* 2012 Nov;17(4):758-770. [doi: [10.1111/j.2044-8287.2012.02070.x](https://doi.org/10.1111/j.2044-8287.2012.02070.x)] [Medline: [22385098](https://pubmed.ncbi.nlm.nih.gov/22385098/)]
69. Neal DT, Wood W, Wu M, Kurlander D. The pull of the past: When do habits persist despite conflict with motives? *Pers Soc Psychol Bull* 2011;37:1428-1437. [doi: [10.1177/0146167211419863](https://doi.org/10.1177/0146167211419863)]
70. Weingarten HP, Elston D. Food cravings in a college population. *Appetite* 1991 Dec;17(3):167-175. [Medline: [1799279](https://pubmed.ncbi.nlm.nih.gov/1799279/)]
71. Livshits G, Malkin I, Williams FM, Hart DJ, Hakim A, Spector TD. Longitudinal study of variation in body mass index in middle-aged UK females. *Age (Dordr)* 2012 Oct;34(5):1285-1294 [FREE Full text] [doi: [10.1007/s11357-011-9299-0](https://doi.org/10.1007/s11357-011-9299-0)] [Medline: [21853263](https://pubmed.ncbi.nlm.nih.gov/21853263/)]
72. Wood C, Conner M, Sandberg T, Godin G, Sheeran P. Why does asking questions change health behaviours? The mediating role of attitude accessibility. *Psychol Health* 2014;29(4):390-404 [FREE Full text] [doi: [10.1080/08870446.2013.858343](https://doi.org/10.1080/08870446.2013.858343)] [Medline: [24245778](https://pubmed.ncbi.nlm.nih.gov/24245778/)]
73. Armitage CJ. Evidence that implementation intentions reduce dietary fat intake: a randomized trial. *Health Psychol* 2004 May;23(3):319-323. [doi: [10.1037/0278-6133.23.3.319](https://doi.org/10.1037/0278-6133.23.3.319)] [Medline: [15099174](https://pubmed.ncbi.nlm.nih.gov/15099174/)]
74. Michie S, Dormandy E, Marteau TM. Increasing screening uptake amongst those intending to be screened: the use of action plans. *Patient Educ Couns* 2004 Nov;55(2):218-222. [doi: [10.1016/j.pec.2003.09.005](https://doi.org/10.1016/j.pec.2003.09.005)] [Medline: [15530757](https://pubmed.ncbi.nlm.nih.gov/15530757/)]
75. Jackson C, Lawton R, Knapp P, Raynor DK, Conner M, Lowe C, et al. Beyond intention: do specific plans increase health behaviours in patients in primary care? A study of fruit and vegetable consumption. *Soc Sci Med* 2005 May;60(10):2383-2391. [doi: [10.1016/j.socscimed.2004.10.014](https://doi.org/10.1016/j.socscimed.2004.10.014)] [Medline: [15748685](https://pubmed.ncbi.nlm.nih.gov/15748685/)]
76. Rutter DR, Steadman L, Quine L. An implementation intentions intervention to increase uptake of mammography. *Ann Behav Med* 2006 Oct;32(2):127-134. [doi: [10.1207/s15324796abm3202_10](https://doi.org/10.1207/s15324796abm3202_10)] [Medline: [16972810](https://pubmed.ncbi.nlm.nih.gov/16972810/)]
77. Skår S, Sniehotta FF, Molloy GJ, Prestwich A, Araújo-Soares V. Do brief online planning interventions increase physical activity amongst university students? A randomised controlled trial. *Psychol Health* 2011 Apr;26(4):399-417. [doi: [10.1080/08870440903456877](https://doi.org/10.1080/08870440903456877)] [Medline: [20830646](https://pubmed.ncbi.nlm.nih.gov/20830646/)]
78. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology* 2006;38:69-119. [doi: [10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)]

79. 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 Jan;9(1):e1 [FREE Full text] [doi: [10.2196/jmir.9.1.e1](https://doi.org/10.2196/jmir.9.1.e1)] [Medline: [17478410](https://pubmed.ncbi.nlm.nih.gov/17478410/)]
80. 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](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]

Abbreviations

BMI: body mass index
FFQ: food frequency questionnaire
HRV: heart rate variability
MET: metabolic equivalent of task
MI: motivational interviewing
RMSSD: root mean square of successive differences
SDRR: standard deviation of RR
SDT: self-determination theory
WHR: waist-to-hip ratio

Edited by G Eysenbach; submitted 14.05.14; peer-reviewed by D O'Connor, P Stumbo; comments to author 16.07.14; revised version received 11.08.14; accepted 16.08.14; published 10.10.14.

Please cite as:

Tapper K, Jiga-Boy G, Maio GR, Haddock G, Lewis M

Development and Preliminary Evaluation of an Internet-Based Healthy Eating Program: Randomized Controlled Trial

J Med Internet Res 2014;16(10):e231

URL: <http://www.jmir.org/2014/10/e231/>

doi: [10.2196/jmir.3534](https://doi.org/10.2196/jmir.3534)

PMID: [25305376](https://pubmed.ncbi.nlm.nih.gov/25305376/)

©Katy Tapper, Gabriela Jiga-Boy, Gregory R Maio, Geoffrey Haddock, Michael Lewis. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 10.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Posttreatment Attrition and Its Predictors, Attrition Bias, and Treatment Efficacy of the Anxiety Online Programs

Ali M AL-Asadi^{1,2}, MA, MSc; Britt Klein^{3,4,5}, DPsych (Clinical); Denny Meyer¹, DBL

¹School of Health Sciences, Swinburne University of Technology, Hawthorn, Australia

²Department of Arts and Education, Grande Prairie Regional College, Grande Prairie, AB, Canada

³DVC-Research & Innovation Portfolio; the School of Health Sciences and; the Collaborative Research Network, Federation University, Ballarat, Australia

⁴National Institute of Mental Health Research, The Australian National University, Canberra, Australia

⁵National eTherapy Centre, School of Health Sciences, Swinburne University of Technology, Hawthorn, Australia

Corresponding Author:

Ali M AL-Asadi, MA, MSc

Department of Arts and Education

Grande Prairie Regional College

10726 - 106 Ave

Grande Prairie, AB, T8V 4C4

Canada

Phone: 1 780 539 2061

Fax: 1 780 538 0960

Email: aalasadi@gprc.ab.ca

Abstract

Background: Although relatively new, the field of e-mental health is becoming more popular with more attention given to researching its various aspects. However, there are many areas that still need further research, especially identifying attrition predictors at various phases of assessment and treatment delivery.

Objective: The present study identified the predictors of posttreatment assessment completers based on 24 pre- and posttreatment demographic and personal variables and 1 treatment variable, their impact on attrition bias, and the efficacy of the 5 fully automated self-help anxiety treatment programs for generalized anxiety disorder (GAD), social anxiety disorder (SAD), panic disorder with or without agoraphobia (PD/A), obsessive-compulsive disorder (OCD), and posttraumatic stress disorder (PTSD).

Methods: A complex algorithm was used to diagnose participants' mental disorders based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision; DSM-IV-TR). Those who received a primary or secondary diagnosis of 1 of 5 anxiety disorders were offered an online 12-week disorder-specific treatment program. A total of 3199 individuals did not formally drop out of the 12-week treatment cycle, whereas 142 individuals formally dropped out. However, only 347 participants who completed their treatment cycle also completed the posttreatment assessment measures. Based on these measures, predictors of attrition were identified and attrition bias was examined. The efficacy of the 5 treatment programs was assessed based on anxiety-specific severity scores and 5 additional treatment outcome measures.

Results: On average, completers of posttreatment assessment measures were more likely to be seeking self-help online programs; have heard about the program from traditional media or from family and friends; were receiving mental health assistance; were more likely to learn best by reading, hearing and doing; had a lower pretreatment Kessler-6 total score; and were older in age. Predicted probabilities resulting from these attrition variables displayed no significant attrition bias using Heckman's method and thus allowing for the use of completer analysis. Six treatment outcome measures (Kessler-6 total score, number of diagnosed disorders, self-confidence in managing mental health issues, quality of life, and the corresponding pre- and posttreatment severity for each program-specific anxiety disorder and for major depressive episode) were used to assess the efficacy of the 5 anxiety treatment programs. Repeated measures MANOVA revealed a significant multivariate time effect for all treatment outcome measures for each treatment program. Follow-up repeated measures ANOVAs revealed significant improvements on all 6 treatment outcome measures for GAD and PTSD, 5 treatment outcome measures were significant for SAD and PD/A, and 4 treatment outcome measures were significant for OCD.

Conclusions: Results identified predictors of posttreatment assessment completers and provided further support for the efficacy of self-help online treatment programs for the 5 anxiety disorders.

Trial Registration: Australian and New Zealand Clinical Trials Registry ACTRN121611000704998; http://www.anzctr.org.au/trial_view.aspx?ID=336143 (Archived by WebCite at <http://www.webcitation.org/618r3wvOG>).

(*J Med Internet Res* 2014;16(10):e232) doi:[10.2196/jmir.3513](https://doi.org/10.2196/jmir.3513)

KEYWORDS

posttreatment attrition; posttreatment predictors; treatment efficacy; online therapy; e-mental health; cognitive behavioral therapy; Internet interventions; fully automated; self-help; Web treatment; generalized anxiety disorder; obsessive compulsive disorder

Introduction

In this age of technological advancement and the increase in peoples' comfort in using the Internet and online resources, online therapy promises to provide an alternative methodology to face-to-face therapy and to be an effective vehicle to deliver treatment to individuals suffering from a variety of psychological disorders. The development and dissemination of e-mental health services have increased at an exponential rate [1]. This relatively new development has been progressing over the past decade and will likely play an important role in reshaping health care over the next decade [2,3].

Various specific types of e-mental health services exist, such as online counseling, mental health information websites, self-guided treatment programs, and online support groups. However, those providing a number of core psychological functions (eg, assessment, referral, treatment) are commonly referred to as online or virtual clinics [4-6]. Here, before accessing an online treatment program, some kind of assessment process typically occurs to screen for mental health problems that are less suitable for online therapy, such as imminent risk of suicide, as well as to help ensure the delivery of appropriate treatment of the client's particular psychological concern [7,8]. Crisis management and/or external referral commonly occur at this stage when the person is at risk / not suitable. The assessment is then generally followed by access to a structured treatment program that often incorporates cognitive behavioral therapy (CBT) techniques because of the established efficacy of this therapeutic modality [9-11]. Once the treatment is completed, a posttreatment assessment is scheduled. Therefore, it is reasonable to consider attrition and its predictors at each 1 of the 3 treatment cycle stages: pretreatment attrition, during treatment attrition, and posttreatment attrition. Examination of attrition and its predictors at each stage would help with early identification of those at risk of dropping out. Being able to do so may result in developing better designs and more targeted interventions to reduce dropouts or noncompletions.

Numerous articles have highlighted the increasing popularity and the rapid growth of online interventions, (eg, [6,8,12-14]) and have discussed the effectiveness of online programs for the treatment of anxiety disorders (eg, [2,6,15-19]). However, many trials are typically small to moderate in sample size (see [16,20]) and include relatively high rates of attrition [12,21,22]; this continues to be a limiting and challenging factor at each stage in the process. Although many papers indirectly provide the specific study dropout rates at the various stages of their study via flowcharts, attrition is generally discussed as a single category (ie, combining pretreatment assessment, during treatment and posttreatment assessment / follow-up attrition

together). Similarly, most of the research published to date includes little or no analysis of predictors of attrition at the different stages in the study process, especially pretreatment attrition predictors or attrition bias and its impact on treatment efficacy. Instead, conservative intention-to-treat analyses, which are likely to attenuate differences [16], have been performed to evaluate the efficacy of online treatments, usually with last observation carried forward (LOCF) imputation for attrition cases [23] or the more advanced method of multiple imputation [24-26]. The scope to which intention-to-treat has been used in online treatment research is not clear, but it is recommended and widely used in most published large-scale studies [20]. However, if it can be shown that attrition bias is unlikely, then completer analysis is likely to be a reasonable and accurate approach. Furthermore, if it can be shown that attrition bias is unlikely and if measures are available on more than 2 occasions, a more accurate maximum likelihood longitudinal analysis would be possible [27,28] and highly recommended [29].

Attrition bias notwithstanding, many reviews and discussion papers firmly support the efficacy of e-mental health programs for multiple psychological concerns, such as anxiety disorders, depression, alcohol and other drug problems, and even when these disorders coexist as comorbid conditions (eg, [8,30]). Andersson et al [31] concluded that existing research supported the efficacy of online programs for the treatment of anxiety disorders, especially when a form of CBT framework was used coupled with some client-therapist contact. Amstadter et al [32] found similar results and concluded that common CBT techniques, such as cognitive restructuring and relapse prevention, can be adapted to online programs with relative ease. More recently, Andersson and Titov [6] concluded that evidence supporting the efficacy of self-help Internet-delivered CBT was strong and consistent.

In summary, online therapy is becoming more popular and more studies are attesting to its utilization and efficacy for a variety of mental health issues and disorders. Most online treatment programs consist of structured modules that include CBT techniques because of its established efficacy. Although several studies have examined predictors of attrition for online therapy, studies looking at the predictors at specific treatment time points (pretreatment, during, posttreatment attrition) are very limited. There is also a lack of studies that include analysis of attrition bias and its impact on treatment efficacy.

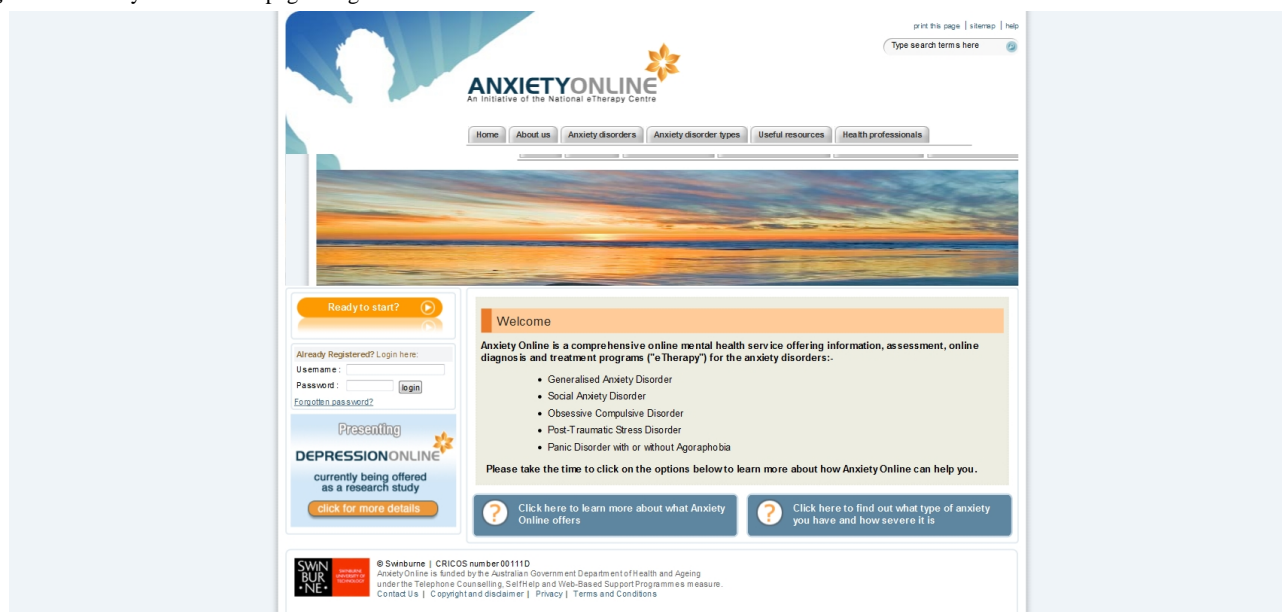
In this study, posttreatment assessment analysis of attrition and its predictors and the treatment efficacy of the Anxiety Online programs were examined. Identifying the characteristics of those who completed the posttreatment assessment measures will assist in intervening early and in devising ways and directing attention to those who do not complete the posttreatment

assessment measures. Anxiety Online is an open-access virtual clinic providing online assessment and diagnosis of 21 mental health disorders defined by the *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision; *DSM-IV-TR*) and self-help and therapist-assisted treatment programs for the 5 anxiety disorders (Figure 1) [33]. The Anxiety Online platform was upgraded in September 2013 and now uses the name Mental Health Online [33].

Previously, we examined pretreatment attrition and during treatment formal withdrawal and their predictors in the Anxiety

Online data [34]. In this study, we reported on 24 demographic and personal variables and 1 treatment variable that were potentially associated with posttreatment attrition. The impact of attrition bias on the treatment outcome measures was assessed by using the data collected on the 5 fully automated self-help online programs from October 2009 to January 2012. When attrition bias was found to be unlikely, a completer analysis of treatment efficacy was performed based on the pre- and posttreatment treatment outcome measures.

Figure 1. Anxiety Online homepage image.



Methods

Procedure

The Anxiety Online platform consists of 4 centers: psychoeducational, assessment, treatment, and health care professional training. The psychoeducational center is a website that provides psychoeducational information about prevalence, symptoms, and treatment of anxiety disorders as well as links to useful websites. The assessment center contains electronic psychological assessment screening system (e-PASS) that consists of a demographic/personal questionnaire and the online diagnostic program (together called *assessment measures* henceforth). As shown in [Multimedia Appendix 1](#), the demographic/personal questionnaire contains a total of 24 demographic and personal variables. After completing the questionnaire, a person then completes the e-PASS that consists of more than 100 diagnostic questions, including the Kessler-6 [35] and items that screen for suicide risk and psychosis (see [36] for details). The treatment center provides and manages the 5 anxiety disorder-specific treatment programs. The training center provides the online therapist training programs and the health care practitioner portal. Individuals can access the Anxiety Online service from anywhere in the world via an Internet connection. People complete the e-PASS if they are interested in psychological assessment function and/or if they are interested in online treatment programs. Based on an individual's response to some of the questions of the e-PASS,

a person may be given a primary diagnosis and/or multiple secondary diagnoses in accordance with *DSM-IV-TR* criteria. Those adults (aged 18 years or older) who receive a primary or secondary diagnosis of panic disorder with or without agoraphobia (PD/A), social anxiety disorder (SAD), posttraumatic stress disorder (PTSD), generalized anxiety disorder (GAD), or obsessive-compulsive disorder (OCD) are offered an online 12-week self-help or therapist-assisted treatment program (the therapist-assisted program is only available to Australian residents). Once participants are enrolled into 1 of the 5 fully automated 12-week self-help treatment programs, they cannot enroll in another online program; however, they can opt out of the treatment program by using the "opt out" option available within the program. Those participants who do not opt out are sent automated emails, with several reminders over a 3-week period following their 12-week treatment cycle, asking them to complete the posttreatment assessment measures. The posttreatment measures are essentially the same as the pretreatment measures. Participants are encouraged to complete the e-PASS annually for 5 years following treatment program cycle completion. Those who want to undertake e-PASS are first required to register and consent to the Anxiety Online terms and conditions [33]. The procedures for collecting and reporting of the Anxiety Online data were approved by the Swinburne University Human Research Ethics Committee. From the time of its launch to the public in October

2009 until January 2012, the e-PASS program has been accessed by 10,745 people.

Treatment Outcome Measures

We identified 6 outcome measures that may be used to indicate successful treatment. The first and second outcome measures were the severity of anxiety disorder--specific symptoms and the severity of major depressive episode (MDE). The disorder-specific severity score is the average of the scores on 6 questions measured on a 8-point Likert scale that assess the level of distress and how much the symptoms of a given disorder interfere in one's life (see [36]). A reduction in the severity score would suggest a positive treatment outcome. The third outcome measure was the number of diagnosed primary and/or secondary disorders. Based on the individual's responses to the e-PASS, each person was given 1 primary disorder and 1 or more secondary disorders if warranted. A reduction in the number of diagnosed disorders (called *number of disorders* henceforth) was indicative of a successful treatment outcome. The fourth outcome measure was the total score on the Kessler-6. The Kessler-6 consists of 6 items measured on a 5-point Likert scale measuring nonspecific psychological distress over the past 30 days. Normative data indicate that 71.7% of the population report low distress scores of 6 to 11, 16.6% of the population report moderate distress scores of 12 to 15, 7.16% of the population report high distress scores of 16 to 19, whereas 2.5% of the population report very high distress scores of 20 to 30 [35,37]. A reduction in the Kessler-6 total score would also suggest a successful treatment outcome. The fifth outcome measure was the individual's self-confidence in managing one's mental health issues. This self-confidence measure is a self-report question measured on a 5-point Likert scale as shown in [Multimedia Appendix 1](#). An increase in reported self-confidence is suggestive of a positive treatment outcome. The final outcome measure was the individual's perceived overall quality of life. The quality-of-life measure is a self-report question measured on a 5-point Likert scale as shown in [Multimedia Appendix 1](#). Similar to self-confidence, an increase in the reported overall quality of life was indicative of a positive treatment outcome.

Participants

As shown in [Figure 2](#), a total of 10,745 individuals completed the pretreatment assessment measures between October 2009 and January 2012. Some of those individuals were people younger than age 18 years (n=202) and some were professionals (n=45) who were exploring the assessment instrument. These 247 individuals were removed from the data leaving 10,498 valid completers of the e-PASS program. In addition, another 249 individuals who did not receive an e-PASS diagnosis of any of the 21 disorders and another 855 individuals who did not receive an e-PASS primary or secondary diagnosis of any of the 5 anxiety disorders were also removed. The removal of these individuals resulted in a sample consisting of 4771 (50.79%) individuals whose primary diagnosis and 4623 (49.21%) individuals whose secondary diagnosis was 1 of the 5 anxiety disorders, for a total of 9394 e-PASS pretreatment completers. All 9394 were offered a treatment program, although it was recommended that those with a primary diagnosis other

than anxiety should seek help elsewhere. Only 3880 individuals accepted and commenced a 12-week online treatment, whereas 5514 individuals did not accept the offer of an online treatment program. At the time of analysis, there were 539 individuals still undergoing treatment; 3199 individuals had not formally withdrawn from their treatment program cycle, whereas 142 individuals formally withdrew during their treatment program cycle.

Unfortunately, of the 3199 individuals, only 383 (11.97%) individuals in total completed the posttreatment assessment measures, whereas 2816 (88.03%) individuals did not complete. Of the 3199, a total of 92 (2.88%) individuals selected the therapist-assisted therapy (36 individuals completed the posttreatment assessment measures and 56 did not complete), whereas 3017 (97.12%) individuals selected 1 of the self-help online programs (347 individuals completed the posttreatment assessment measures and 2760 did not complete). To keep the focus on self-help and exclude any therapist intervention, for the purpose of this analysis, we only considered the 347 individuals who selected the self-help online treatment programs and completed the posttreatment assessment measures in comparison with the 2760 individuals who also selected self-help online treatment programs but did not complete the posttreatment assessment measures. The distribution of individuals who enrolled in the 5 online treatment programs and whether they completed or did not complete the posttreatment assessment measures is shown in [Table 1](#).

The self-help online noncompleter group consisted of 860 males whose age ranged between 18 and 78 years with a mean of 37.74 (SD 12.05) years and 1900 females whose age ranged between 18 and 81 years with a mean of 35.11 (SD 11.57) years. The males group (n=860) consisted of 614 (71.4%) males who reported living in metropolitan areas, 176 (20.5%) males in regional areas, 63 (7.3%) males in rural areas, and 7 (0.8%) males who reported living in remote areas. The females group (n=1900) consisted of 1204 (63.37%) females who reported living in metropolitan areas, 447 (23.53%) females in regional areas, 230 (12.11%) females in rural areas, and 19 (0.01%) females who reported living in remote areas.

The self-help online completer group consisted of 117 males whose age ranged between 19 and 75 years with a mean of 42.19 (SD 12.72) years and 230 females whose age ranged between 18 and 75 years with a mean of 40.34 (SD 12.68) years. The males group (n=117) consisted of 76 (65.0%) males who reported living in metropolitan areas, 31 (26.5%) males in regional areas, 9 (7.7%) males in rural areas, and 1 (0.9%) male who reported living in remote areas. The females group (n=230) consisted of 144 (62.6%) females who reported living in metropolitan areas, 59 (25.7%) females in regional areas, 26 (11.3%) females in rural areas, and 1 (0.4%) female who reported living in remote areas.

The first part of the analysis investigated the differences between the posttreatment assessment completers and noncompleters for the 5 anxiety treatment programs for the Kessler-6 total score and the demographic and personal variables shown in [Multimedia Appendix 1](#). Posttreatment attrition is defined as the ratio of posttreatment assessment noncompleters to the total

number of participants, and it is this measure and its predictors that are the focus of this study. Because of the large number of individuals who chose not to complete the posttreatment assessment measures, the second part of the analysis investigated the potential impact of posttreatment attrition bias on the treatment outcome measures of the 5 fully self-help online treatment programs. By using G*Power 3.1 [38], a minimum

of 34 participants in each of the 5 anxiety treatment programs were needed to achieve a power of 80% ($\alpha=.05$) to detect a moderate effect size of 0.5 on treatment outcomes. The number of participants in the 5 treatment programs ranged from 36 to 134 participants, so the interpretation of these results should be considered reliable.

Figure 2. Recruitment and enrollment rate throughout the process.

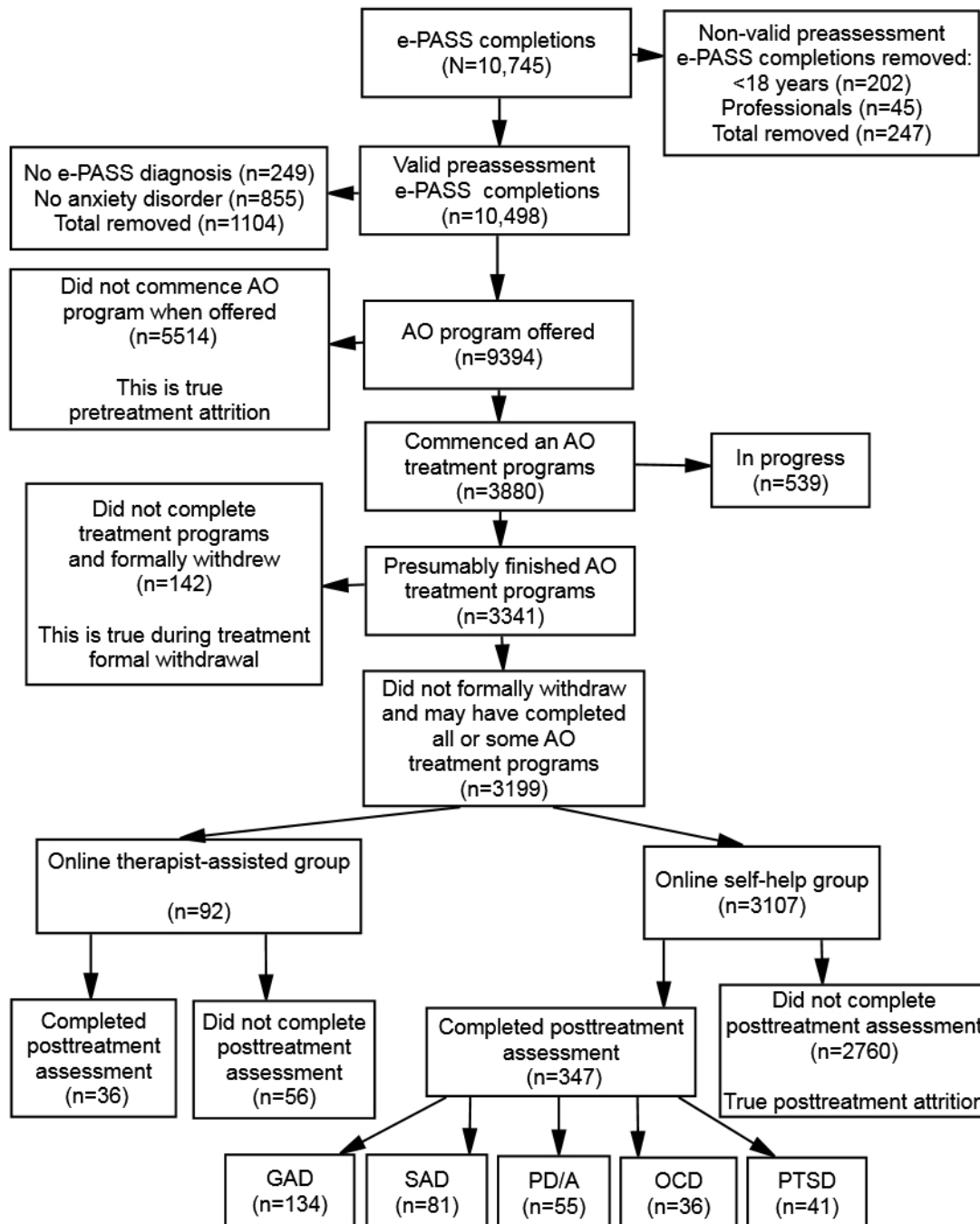


Table 1. Number of individuals enrolled in the self-help online programs that completed or did not complete the posttreatment assessment measures.

Anxiety online programs	Self-help, n (%)		Total
	Completed posttreatment assessment	Did not complete posttreatment assessment	
GAD	134 (11.57)	1024 (88.43)	1158
SAD	81 (9.56)	766 (90.44)	847
PD/A	55 (9.95)	498 (90.05)	553
PTSD	41 (13.76)	257 (86.24)	298
OCD	36 (14.34)	215 (85.66)	251
Total	347 (11.17)	2760 (88.83)	3107

Analysis

The initial univariate analysis used chi-square tests of association to determine which of the pretreatment assessment demographic and personal variables had a significant relationship with posttreatment attrition. A multivariate analysis was then used to confirm the univariate results. Multivariate binary logistic regression analysis with a forward selection approach was performed to identify the significant predictor variables. The final model was evaluated using a Hosmer-Lemeshow test.

To assess attrition bias, Heckman's 3-step method was used [39]. The first step involves using the attrition demographic and personal predictor variables in a binary logistic regression to predict the probability that each participant will complete the posttreatment assessment measures. The second step involves using this predicted probability to compute the Mills ratio for each participant using the ratio of the normal probability and cumulative distribution function for the residuals. The third step involves determining the effect of (predicted) attrition risk on the change in treatment outcome measures between pre- and posttreatment. This is accomplished by constructing a multivariate general linear model based on the difference between pre- and posttreatment treatment outcome measures and treatment programs with the Mills ratio entered as a covariate. A nonsignificant Mills ratio effect and a nonsignificant Mills ratio by programs interaction effect are indicative of no attrition bias. That is, noncompleters of posttreatment assessment measures would have given similar answers to those given by completers.

Finally, repeated measures MANOVA followed by repeated measures ANOVAs for the treatment outcome measures (anxiety-specific severity rating, MDE severity rating, Kessler-6 total score, number of disorders, self-confidence, and quality

of life) were used to evaluate the 5 self-help online treatment programs separately. The data supported the assumptions of normality and homogeneity. The Cohen's *d* effect size classification scheme (0.20 for small effect, 0.50 for moderate effect, and 0.80 for large effect) was used [40]. SPSS version 20 (IBM Corp, Armonk, NY, USA) was used to produce all results.

Results

Posttreatment Attrition (Profile of Posttreatment Assessment Completers)

As shown in Table 1, most participants (2760/3107, 88.83%) enrolled in 1 of the 5 self-help online treatment programs, did not formally opt out, but did not complete the posttreatment assessment measures. Only 347 participants completed the posttreatment assessment measures yielding a posttreatment assessment completion rate of 11.17% (347/3107) and a posttreatment attrition rate of 88.83% (2760/3107).

As shown in Table 2, the chi-square tests of association and *F* tests showed 9 variables that were significantly associated with completing the posttreatment assessment measures. Those who completed the posttreatment assessment measures tended to differ from those who did not complete posttreatment assessment measures in several ways. On average, it was more likely that the completers of posttreatment assessment measures had heard about the program through traditional media rather than the Internet, were seeking online assistance with the primary goal of finding a self-help program; were willing to provide consumer feedback, were receiving mental health assistance, were nonsmokers, rated their self-confidence as "good," said that they learned best by reading, had a slightly lower pretreatment Kessler-6 total score, and were older in age.

Table 2. Predictor analysis for attrition categories for posttreatment assessment completers and noncompleters for online self-help group (N=3107).

Variables	Attrition categories		Test of association		
	Completers (n=347)	Noncompleters (n=2760)	χ^2 (df)	F _{1, 3105}	P
How did you hear about us?, n (%)			25.6 (4)		.001
Internet	108 (31.1)	1205 (43.66)			
Health professional	58 (16.7)	407 (14.75)			
Friend/family	20 (5.8)	198 (7.17)			
Traditional media	104 (20.0)	623 (22.57)			
Other	57 (16.4)	327 (11.85)			
Reason for seeking online assistance, n (%)			13.0 (1)		.001
To complete 1 of the self-help programs	231 (66.6)	1557 (56.41)			
Provide consumer feedback (yes), n (%)	182 (52.5)	1285 (46.56)	8.6 (1)		.01
Currently receiving mental health assistance (yes), n (%)	146 (42.1)	986 (35.72)	5.4 (1)		.02
Do you smoke? (yes), n (%)	41 (11.8)	464 (16.81)	5.7 (1)		.02
Self-confidence, n (%)			13.3 (4)		.01
Very poor	15 (4.3)	193 (6.99)			
Poor	76 (21.9)	716 (25.94)			
Neither	132 (38.0)	1049 (38.01)			
Good	114 (32.9)	693 (25.11)			
Very good	10 (2.9)	109 (3.95)			
How do you best learn?, n (%)			13.6 (3)		.004
Hearing	25 (7.2)	151 (5.47)			
Reading	125 (36.0)	811 (29.38)			
Looking and watching	43 (12.4)	522 (18.91)			
Doing	154 (44.4)	1276 (46.23)			
Pre-Kessler-6 (total score), mean (SD)	15.97 (4.90)	16.99 (4.80)		13.73	.001
Age (years), mean (SD)	40.97 (12.71)	35.93 (11.79)		55.29	.001

As shown in [Table 3](#), the final binary logistic regression with forward selection of predictors for posttreatment attrition contained 6 significant predictors. The Hosmer-Lemeshow goodness-of-fit test indicated an adequate model fit. When statistically controlling for the other variables in the model, we found significant odds ratios for the following predictors: how participants first heard about the Anxiety Online program, reason for registering, age, whether currently receiving assistance for mental health concerns, best method of learning, and pre-Kessler-6 total score. We should note that the same predictors were also found to be significantly associated with posttreatment attrition using chi-square tests as shown in [Table 2](#).

The expected odds for completing the posttreatment assessment measures in order of significance were as follows: 3% increase

in likelihood for each year increase in age; 3% reduction in likelihood for each additional point an individual scored on the Kessler-6 total score; 1.76 and 1.42 times higher for those who heard about the Anxiety Online from the traditional media (TV, radio, magazine, newspaper) and from friends or family members, respectively, relative to other sources (eg, brochure, mail-out, newsletter, e-bulletin, lecture, conference, support group, through work, Facebook); 1.42 times higher for individuals who gave “seeking to use 1 of the self-help online programs” as a reason for joining the program relative to all other reasons; 1.94 and 1.76 times higher for those indicating that they learn best by hearing or reading, respectively, relative to those who said they learn best by looking or watching; and 1.40 times higher for those who reported that they were receiving mental health assistance relative to those who were not receiving mental health assistance.

Table 3. Binary logistic regression model for posttreatment assessment attrition.

Variables	Wald (df)	<i>P</i>	OR (95% CI)
Heard (reference group: other sources)	11.79 (4)	.02	
Internet	3.30 (1)	.07	1.38 (0.98-1.96)
Health professional	0.13 (1)	.72	1.10 (0.66-1.82)
Friend or family	5.06 (1)	.02	1.42 (1.05-1.92)
Traditional media	9.92 (1)	.002	1.76 (1.24-2.49)
Reason (online self-help)(reference group: other reasons)	8.20 (1)	.004	1.42 (1.12-1.81)
Age	33.93 (1)	.001	1.03 (1.02-1.04)
Currently receiving mental health assistance (reference group: none)	7.33 (1)	.007	1.40 (1.10-1.78)
Learning (reference group: looking/watching)	11.09 (3)	.01	
By hearing	5.95 (1)	.02	1.94 (1.14-3.32)
By reading	9.03 (1)	.003	1.76 (1.22-2.55)
By doing	3.37 (1)	.07	1.40 (0.98-2.00)
Pre-Kessler-6 total score	6.64 (1)	.01	0.97 (0.94-0.99)
Constant	100.15 (1)	.001	0.03

Attrition Bias and Evaluation of Treatment Outcomes

Six personal and demographic variables were found to be significantly associated with posttreatment attrition: age, Kessler-6 total score, “how did you first hear about the Anxiety Online?,” reason for registering, “how do you best learn?,” and whether the person was currently receiving mental health assistance. These 6 variables were used in a binary logistic regression to predict the attrition category for all participants. A probability estimate of completing the posttreatment measures was calculated for each participant who actually completed the posttreatment assessment measures. The Mills ratios for all 347 completers of posttreatment assessment measures were calculated using Heckman’s method [39].

Next, the difference between the pretreatment score and the posttreatment score on the 10 treatment outcome measures (5 severity scores for each anxiety disorder, MDE severity score, Kessler-6 total score, number of disorders, self-confidence, and quality of life) for each participant was calculated. To analyze attrition bias for the 347 clients who selected 1 of the online self-help anxiety treatment programs, a MANOVA analysis was carried out to compare the improvement in the 10 treatment outcome measures (represented by the differences in scores of these outcome measures at pre- and posttreatment) for the 5

treatment programs with Mills ratio included as a covariate. Results revealed that the Mills ratio had no significant effect ($F_{10, 328}=0.550, P=.85$) and that there was no significant interaction effect between the Mills ratio and the programs ($F_{40, 1246}=0.683, P=.93$). Therefore, these results suggested that there was unlikely to be any attrition bias for all 5 fully automated anxiety treatment programs, allowing completer analysis to be used.

Analysis of Treatment Outcomes

Overview

A repeated measures MANOVA for each fully automated self-help online treatment program was carried out with the 6 treatment outcome measures (Kessler-6 total score, number of disorders, self-confidence, quality of life, MDE severity, and the anxiety-specific severity measure) followed up by repeated measures ANOVAs with analysis of effect size using Cohen’s *d*. To arrive at more conservative and more accurate values for the effect size, the correlations (*r*) between the pre- and posttreatment outcome measures (reported in Table 4) were ignored in the calculation of Cohen’s *d* and its 95% CI [41]. Results of these analyses, means, standard deviations, and other parameters are displayed in Table 4.

Table 4. Summary of means, standard deviations, correlations for pre- and posttreatment results, *F* values, *P* value, and Cohen's *d* and its 95% CI for treatment outcome measures.

Treatment outcome measures	Pretreatment, mean (SD)	Posttreatment, mean (SD)	<i>F</i> (<i>df</i>)	<i>P</i>	<i>r</i>	Cohen's <i>d</i> (95% CI)
GAD (n=134)						
GAD severity	3.28 (1.54)	2.00 (1.67)	85.35 (1,133)	.001	.50	0.80 (0.54, 1.08)
MDE severity	2.07 (2.07)	1.23 (1.89)	28.36 (1,133)	.001	.57	0.42 (0.07, 0.74)
Kessler-6	16.59 (4.57)	13.62 (4.30)	99.89 (1,133)	.001	.70	0.67 (–0.10, 1.40)
# of disorders	4.31 (1.98)	3.46 (2.18)	34.40 (1,133)	.001	.68	0.41 (0.07, 0.78)
Self-confidence	3.07 (0.88)	3.63 (0.82)	49.40 (1,133)	.001	.41	–0.66 (–0.81, –0.52)
Quality of life	3.43 (0.82)	3.66 (0.82)	10.86 (1,133)	.001	.51	–0.28 (–0.42, –0.14)
SAD (n=81)						
SAD severity	3.12 (1.63)	1.95 (1.87)	37.17 (1,80)	.001	.52	0.67 (0.31, 1.08)
MDE severity	1.60 (1.99)	1.46 (2.17)	0.29 (1,80)	.59	.37	0.07 (–0.37, 0.54)
Kessler-6	15.59 (4.97)	13.72 (4.86)	12.34 (1,80)	.001	.52	0.38 (–0.70, 1.44)
# of disorders	4.46 (2.16)	3.60 (2.15)	17.53 (1,80)	.001	.64	0.40 (–0.07, 0.87)
Self-confidence	3.02 (0.88)	3.47 (0.92)	12.55 (1,80)	.001	.22	–0.50 (–0.69, –0.30)
Quality of life	3.32 (0.91)	3.54 (0.90)	5.71 (1,80)	.02	.57	–0.24 (–0.44, –0.05)
OCD (n=36)						
OCD severity	2.81 (1.87)	1.97 (2.37)	5.25 (1,35)	.03	.48	0.40 (–0.21, 1.17)
MDE severity	1.76 (2.41)	1.04 (1.95)	6.50 (1,35)	.02	.72	0.33 (–0.46, 0.97)
Kessler-6	14.42 (5.60)	13.28 (5.62)	2.67 (1,35)	.11	.72	0.20 (–1.63, 2.04)
# of disorders	4.33 (2.15)	3.03 (2.25)	15.16 (1,35)	.001	.58	0.59 (–0.11, 1.33)
Self-confidence	3.25 (0.91)	3.81 (0.82)	12.59 (1,35)	.001	.41	–0.65 (–0.94, –0.38)
Quality of life	3.58 (0.87)	3.75 (1.03)	2.69 (1,35)	.11	.81	–0.18 (–0.46, –0.16)
PD/A (n=55)						
PD/A severity	3.19 (1.87)	1.51 (2.14)	35.41 (1,54)	.001	.46	0.84 (0.34, 1.40)
MDE severity	1.75 (2.05)	1.15 (2.01)	7.41 (1,54)	.009	.67	0.30 (–0.25, 0.83)
Kessler-6	14.89 (4.69)	12.51 (4.69)	22.86 (1,54)	.001	.69	0.51 (–0.73, 1.75)
# of disorders	4.62 (2.16)	3.44 (2.52)	14.91 (1,54)	.001	.54	0.50 (–0.07, 1.17)
Self-confidence	3.05 (1.03)	3.55 (0.90)	18.01 (1,54)	.001	.61	–0.52 (–0.79, –0.28)
Quality of life	3.56 (1.15)	3.67 (0.94)	1.29 (1,54)	.26	.79	–0.11 (–0.41, 0.14)
PTSD (n=41)						
PTSD severity	3.01 (1.74)	2.09 (1.99)	7.21 (1,40)	.01	.32	0.50 (–0.04, 1.10)
MDE severity	2.42 (2.16)	1.65 (2.10)	5.64 (1,40)	.02	.53	0.36 (–0.30, 1.00)
Kessler-6	17.51 (4.93)	14.73 (5.23)	9.19 (1,40)	.004	.33	0.55 (–0.96, 2.15)
# of disorders	5.39 (2.97)	4.32 (3.07)	8.63 (1,40)	.005	.70	0.35 (–0.55, 1.29)
Self-confidence	3.12 (0.95)	3.76 (0.89)	22.35 (1,40)	.001	.57	–0.70 (–0.99, –0.42)
Quality of life	3.10 (0.83)	3.49 (0.90)	15.85 (1,40)	.001	.74	–0.45 (–0.70, –0.18)

Generalized Anxiety Disorder Treatment Program

A significant multivariate time effect was found for the GAD program ($F_{6, 128}=26.85, P<.001$). Subsequent repeated measures ANOVAs showed significant improvements on all 6 treatment outcome measures. GAD severity produced a large effect size, Kessler-6 and self-confidence produced moderate effect sizes,

and MDE severity, number of disorders, and quality of life produced small effect sizes.

Posttraumatic Stress Disorder Treatment Program

A significant multivariate time effect was found for the PTSD program ($F_{6, 35}=4.45, P=.002$). Subsequent repeated measures ANOVAs showed significant improvements on all 6 treatment

outcome measures. Self-confidence, Kessler-6, and PTSD severity produced moderate effect sizes, whereas quality of life, MDE severity, and number of disorders produced small effect sizes.

Social Anxiety Disorder Treatment Program

A significant multivariate time effect was found for the SAD program ($F_{6,75}=8.36, P<.001$). Subsequent repeated measures ANOVAs showed significant improvements on 5 of the 6 treatment outcome measures. SAD severity and self-confidence produced moderate effect sizes, whereas Kessler-6, number of disorders, and quality of life produced small effect sizes. MDE severity produced a very small and nonsignificant effect size.

Panic Disorder With or Without Agoraphobia Treatment Program

A significant multivariate time effect was found for the PD/A program ($F_{6,49}=8.89, P<.001$). Subsequent repeated measures ANOVAs showed significant improvements on 5 of the 6 treatment outcome measures. PD/A severity produced a large effect size, whereas Kessler-6, self-confidence, number of disorders, and MDE severity produced moderate effect sizes. Quality of life produced a small nonsignificant effect size.

Obsessive-Compulsive Disorder Treatment Program

A significant multivariate time effect was found for the OCD program ($F_{6,30}=4.18, P=.004$). Subsequent repeated measures ANOVAs showed significant improvements on 4 of the 6 treatment outcome measures. Self-confidence and number of disorders produced moderate effect sizes, whereas OCD severity and MDE severity produced small but significant effect sizes. Kessler-6 and quality of life produced small nonsignificant effect sizes.

Discussion

The purpose of this study was to examine posttreatment assessment attrition and its predictors, and to assess the potential for attrition bias and its impact on treatment outcome measures for the Anxiety Online self-help programs. The posttreatment assessment attrition rate for the self-help programs was found to be 89%. This is a large posttreatment assessment attrition rate compared with therapist-assisted randomized controlled trials of online treatment (eg, 13% [42], 6.4% [43], and 0% [44]). However, our attrition rate compares favorably with attrition rates reported by similar open-access fully automated treatment programs (eg, [21]).

The e-PASS program collected data on 24 demographic and personal variables and 1 measure of psychological distress, the Kessler-6. Chi-square tests of association and binary logistic regression were used to relate these variables to posttreatment assessment attrition. Results revealed that the likelihood of completing posttreatment assessment measures declined for participants with a greater Kessler-6 total score and increased for older participants, participants who heard about the program through the traditional media and from family and friends, those who were looking to complete a self-help online program, participants receiving assistance for mental health concerns, and for participants who reported learning best by reading,

hearing, and doing rather than looking and watching. Those that joined the program because they wanted to receive online therapy were more likely to complete the posttreatment assessment measures potentially because of greater motivation and commitment to the program and interest around their treatment outcome. Participants who learn best by reading, hearing, and doing would likely be more involved in their learning than those who learn more passively by looking and watching. This difference in the reported learning style between being actively or passively involved may explain why those reporting the former style were more likely than those reporting the latter style to complete the posttreatment assessment measures. Participants who were receiving mental health services were likely to be more invested and actively engaged in managing their mental health and, therefore, they were more likely to show the tendency to complete the posttreatment assessment measures.

Interestingly, older participants were more likely to complete the posttreatment assessment measures, perhaps because with age comes a greater sense of commitment to the task. The age of our sample ranged between 18 and 78 years and because age was not a significant predictor of pretreatment attrition and formal withdrawal during treatment attrition, age was not a discriminatory factor [34]. This is especially encouraging because online programs are often thought to be of greater interest to younger cohorts [45], whereas these results suggest online programs are applicable across the age spectrum.

Similar to traditional face-to-face treatment programs, online programs ideally start with an assessment of the issues, then move to treatment of these issues, and then proceed to assessment of the impact of treatment on these issues. These 3 phases (pretreatment assessment, treatment, and posttreatment assessment) make up the standard design for any treatment program; therefore, attrition and its predictors at these different stages should be examined separately and not viewed as a single category. At each phase, there will be those who will start but not finish. Therefore, it is important to assess not only attrition at each of these 3 phases but also the predictors of attrition. In AL-Asadi et al [34], the predictors of pretreatment attrition and formal withdrawal during treatment were identified. In this study, we identified the predictors of posttreatment attrition. If we were to consider the findings of these 2 studies combined, we discover that there was no single predictor that is present in all phases, that some variables were predictors in only 1 phase, and that only some variables were predictors in 2 phases. For example, we found those who heard about the program from family/friends and traditional media sources, those who were seeking online self-help, those who learn best by reading, and those who had lower levels of psychological distress were more likely to accept and commence treatment and more likely to complete the posttreatment assessment measures. On the other hand, those who were concerned about anxiety, those who reported poor quality of life, and those who were prepared to make or in the process of making changes were more likely to accept and commence treatment and more likely not to formally withdraw from treatment.

By using Heckman's method [39], attrition bias was assessed to be nonsignificant and consequently it is reasonable to

conclude that those who did not complete the posttreatment assessment measures would have responded in the same way as those who completed the posttreatment assessment measures. The nonsignificant attrition bias allowed the use of completer analysis to assess the impact of the treatment programs in place of overconservative intention-to-treat analyses. Analyses of the treatment outcomes revealed that for all 5 treatment programs (GAD, SAD, PD/A, OCD, and PTSD) there was a significant effect in reducing the diagnostic anxiety-specific severity reported by participants, reducing the total number of diagnosed disorders, and increasing the reported self-confidence in dealing with ones' mental health issues.

The MDE severity scores significantly decreased for participants in 4 of the treatment programs with the SAD group showing a small nonsignificant improvement. This finding that the treatment for anxiety disorder produced not only significant reduction in the severity of the anxiety-specific symptoms but also in the severity of symptoms of depression is indicative of the efficacy of online treatment to provide transdiagnostic treatment. These results are consistent with the conclusions of Andersson and Titov [6] and Johansson et al [46].

Psychological distress, as measured by the Kessler-6 total score, significantly decreased for participants in 4 of the treatment programs with the OCD group showing a small nonsignificant improvement. Similarly, the quality-of-life rating significantly improved for participants in 3 of the treatment programs with the OCD and the PD/A groups showing nonsignificant improvement. Overall, these results support the efficacy of online treatment of the 5 anxiety disorders. Cohen's *d* within-group treatment effect sizes ranged from 0.40 to 0.84 for the relevant anxiety-specific severity scores. These results are consistent with the typical range of 0.4-0.7 reported for other self-help online programs [36,47,48].

These results suggest that efficacious fully automated self-help online treatment programs for a variety of anxiety disorders can be delivered to anyone with an Internet connection, anywhere, at any time. This increase in accessibility to treatment should make it easier for those whose mobility is restricted, those who feel uncomfortable being seen in a local mental health clinic, those who do not have local resources, and those who are unable to adhere to regular appointments to access mental health treatment. Online programs may have other potential advantages. For example, the potential to reach large and/or rural populations at a fraction of the cost associated with face-to-face therapy, and the privacy and anonymity of accessing therapy in one's own home reduces the cost as well as the stigmatization [49-51] associated with accessing face-to-face services. Two literature reviews by Musiat and Tarrrier [1] and by Lal and Adair [3] agreed that online programs were cost effective, but disagreed on geographical and time flexibility. Musiat and Tarrrier [1] concluded that there was limited reporting around the advantages of geographical, time flexibility, and stigma, whereas Lal and Adair [3] concluded that geographical flexibility, timing and convenience, and anonymity were the strengths and benefits associated with e-mental health programs. These differing opinions highlight the need for more research. Another added advantage is that online therapy makes it easy to provide people with several means of presentation of educational material, such

as written, video, or audio formats, which can facilitate matching learning methods to preferred learning styles/preferences [3,32]. The mode of online therapy offers a promising and cost-effective alternative to face-to-face therapy [36], especially when considering the underengagement of the public with face-to-face services for mental health concerns [52-54]. This percentage becomes even lower for those who are living far from major urban centers and those who reside in rural communities where concerns over stigma is heightened and access and resources are limited [55,56]. In sum, although evidence is not consistent across all benefits, online therapy remains a potentially viable mode of service delivery that may provide both time- and cost-effective intervention and address obstacles associated with traditional therapy, such as stigma, travel issues, and lengthy waitlists to see a therapist. Therefore, online intervention may provide those who otherwise could not or would not seek out psychological treatment with the opportunity to avoid several limiting obstacles and receive effective therapy.

However, there are 4 major limitations that should be noted. Firstly, Anxiety Online platform (now Mental Health Online) is a cost-free system open to anyone in the world with Internet access. The design of this system does not require a control group; thus, it is difficult to make any conclusion regarding causal relationships between the treatment programs and improvements. Moreover, the lack of a control group and the high rate of posttreatment attrition make any conclusion about the efficacy of this online therapy preliminary. Secondly, the e-PASS uses online assessment procedures exclusively that rely on self-report to determine the diagnoses of participants. The use of automated online assessment for the purpose of assigning diagnoses is a limitation of this study in itself because the reliability of online diagnostic assessment tools has been questioned [6].

Thirdly, a single study that examined the psychometric properties of the e-PASS concluded that the treatment outcome measures have high test-retest reliability and reasonable convergent validity (D Nguyen, unpublished PhD thesis, Swinburne University, 2013). However, the small sample size and some disagreement with structured clinical interviews in terms of the severity levels required for a clinical diagnosis suggest that further validation studies with large sample sizes are needed. Consequently, more validation studies based on the newly released *DSM-5* criteria must be conducted.

Fourthly, the use of completer analysis may overestimate the effectiveness of the treatment programs when attrition bias is suspected. However, in this case, attrition bias was found to be nonsignificant suggesting that the results accurately reflect the true effectiveness of the treatment programs. The use of the more conservative Cohen's *d* values that do not take correlations into account [41] showed that effect sizes were reasonable despite the use of a conservative effect size measure.

As for the Anxiety Online platform, the high posttreatment attrition rate was a weakness of this platform. It appears that sending several automated email reminders over a 3-week period following the 12-week treatment cycle may be a relatively ineffective way to encourage sufficiently large numbers of people to complete the posttreatment measures. However, we

should acknowledge that, although still high, the multiple reminder email reminders may be 1 of the reasons why the Anxiety Online posttreatment attrition rate is slightly lower than other fully automated self-help open-access systems. This is confirmed by the higher posttreatment completion rates for Anxiety Online therapist-assisted program versions (36/92, 39%) involving a weekly email from a trained therapist.

Having participants complete the posttreatment measures is certainly a challenging task. This is probably exacerbated given the participants have already undertaken the assessment measures before treatment and, therefore, they know how demanding the posttreatment assessment will be. Telephone calls after the multiple email reminders may prove useful in further reminding participants, although this would impact on cost and also detracts from the fully automated nature of the system. Alternatively, motivation to complete posttreatment assessment measures may be increased by educating participants on the importance of completing the posttreatment assessment measures to allow the improvement of the treatment programs for future participants. In addition, asking participants to enter

into a “behavioral contract” beyond the terms and conditions might be valuable (eg, pledge commitment and completion of modules and posttreatment assessment measures before they can commence).

Research on e-mental health has been taking place over the past decade or so examining its efficacy with a number of different disorders. However, it is important to continue to investigate a broader range of mental health problems, other therapeutic modalities besides CBT, and the issues related to geographic and time flexibility, stigma, and specific populations—especially older adults. Furthermore, and in view of the high attrition rates, especially those with open-access fully automated self-help online programs, it is recommended that when establishing treatment efficacy, researchers should consider examining the question of attrition bias. If attrition bias is found to be nonsignificant, completer analysis or maximum likelihood longitudinal methods should be used to assess treatment accuracy rather than the overly conservative intention-to-treat analyses.

Acknowledgments

The authors would like to thank the Australian Government Department of Health and Ageing who provided the funding for the development of the Anxiety Online service. e-PASS and the Anxiety Online e-therapy treatment programs are available to the international public at no cost (fully automated self-help). The therapist-assisted version is open only to Australian residents at a minimal cost. The funders of Anxiety Online (Australian Government Department of Health and Ageing) had no other involvement in this study or report.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Self Report Online Questionnaire.

[[PDF File \(Adobe PDF File\), 36KB - jmir_v16i10e232_app1.pdf](#)]

References

1. Musiat P, Tarrrier N. Collateral outcomes in e-mental health: a systematic review of the evidence for added benefits of computerized cognitive behavior therapy interventions for mental health. *Psychol Med* 2014 Feb 19:1-14. [doi: [10.1017/S0033291714000245](https://doi.org/10.1017/S0033291714000245)] [Medline: [25065947](https://pubmed.ncbi.nlm.nih.gov/25065947/)]
2. Klein B. Australian Psychological Society. 2010. E-Interventions and psychology: time to log on URL: <http://www.psychology.org.au/publications/inpsych/2010/feb/klein/> [accessed 2014-09-28] [WebCite Cache ID 6SwH0A1AB]
3. Lal S, Adair CE. E-mental health: a rapid review of the literature. *Psychiatr Serv* 2014 Jan 1;65(1):24-32. [doi: [10.1176/appi.ps.201300009](https://doi.org/10.1176/appi.ps.201300009)] [Medline: [24081188](https://pubmed.ncbi.nlm.nih.gov/24081188/)]
4. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. *Ann Behav Med* 2009 Aug;38(1):4-17. [doi: [10.1007/s12160-009-9130-7](https://doi.org/10.1007/s12160-009-9130-7)] [Medline: [19787305](https://pubmed.ncbi.nlm.nih.gov/19787305/)]
5. Christensen H, Murray K, Clear AL, Bennett K, Bennett A, Griffiths KM. Beacon: a web portal to high-quality mental health websites for use by health professionals and the public. *Med J Aust* 2010 Jun 7;192(11 Suppl):S40-S44. [Medline: [20528708](https://pubmed.ncbi.nlm.nih.gov/20528708/)]
6. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry* 2014 Feb;13(1):4-11 [FREE Full text] [doi: [10.1002/wps.20083](https://doi.org/10.1002/wps.20083)] [Medline: [24497236](https://pubmed.ncbi.nlm.nih.gov/24497236/)]
7. Abbott J, Klein B, Ciechomski L. Best practices in online therapy. *Journal of Technology in Human Services* 2008 Jul 03;26(2-4):360-375. [doi: [10.1080/15228830802097257](https://doi.org/10.1080/15228830802097257)]
8. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting internet intervention research. *Cogn Behav Ther* 2011;40(2):82-97. [doi: [10.1080/16506073.2011.573807](https://doi.org/10.1080/16506073.2011.573807)] [Medline: [25155812](https://pubmed.ncbi.nlm.nih.gov/25155812/)]

9. Butler AC, Chapman JE, Forman EM, Beck AT. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clin Psychol Rev* 2006 Jan;26(1):17-31. [doi: [10.1016/j.cpr.2005.07.003](https://doi.org/10.1016/j.cpr.2005.07.003)] [Medline: [16199119](https://pubmed.ncbi.nlm.nih.gov/16199119/)]
10. Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. *Annu Rev Psychol* 2001;52:685-716 [FREE Full text] [doi: [10.1146/annurev.psych.52.1.685](https://doi.org/10.1146/annurev.psych.52.1.685)] [Medline: [11148322](https://pubmed.ncbi.nlm.nih.gov/11148322/)]
11. Tolin DF. Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clin Psychol Rev* 2010 Aug;30(6):710-720. [doi: [10.1016/j.cpr.2010.05.003](https://doi.org/10.1016/j.cpr.2010.05.003)] [Medline: [20547435](https://pubmed.ncbi.nlm.nih.gov/20547435/)]
12. Christensen H, Griffiths K, Groves C, Korten A. Free range users and one hit wonders: community users of an Internet-based cognitive behaviour therapy program. *Aust N Z J Psychiatry* 2006 Jan;40(1):59-62. [doi: [10.1111/j.1440-1614.2006.01743.x](https://doi.org/10.1111/j.1440-1614.2006.01743.x)] [Medline: [16403040](https://pubmed.ncbi.nlm.nih.gov/16403040/)]
13. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. *J Med Internet Res* 2011;13(3):e52 [FREE Full text] [doi: [10.2196/jmir.1772](https://doi.org/10.2196/jmir.1772)] [Medline: [21821503](https://pubmed.ncbi.nlm.nih.gov/21821503/)]
14. Schmidt U, Wykes T. E-mental health - a land of unlimited possibilities. *J Ment Health* 2012 Aug;21(4):327-331. [doi: [10.3109/09638237.2012.705930](https://doi.org/10.3109/09638237.2012.705930)] [Medline: [22823092](https://pubmed.ncbi.nlm.nih.gov/22823092/)]
15. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
16. Arberg FK, Linton SJ, Hultcrantz M, Heintz E, Jonsson U. Internet-delivered psychological treatments for mood and anxiety disorders: a systematic review of their efficacy, safety, and cost-effectiveness. *PLoS One* 2014;9(5):e98118 [FREE Full text] [doi: [10.1371/journal.pone.0098118](https://doi.org/10.1371/journal.pone.0098118)] [Medline: [24844847](https://pubmed.ncbi.nlm.nih.gov/24844847/)]
17. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: a randomized controlled trial comparing unguided with two types of guided self-help. *Behav Res Ther* 2011 Mar;49(3):158-169. [doi: [10.1016/j.brat.2010.12.007](https://doi.org/10.1016/j.brat.2010.12.007)] [Medline: [21255767](https://pubmed.ncbi.nlm.nih.gov/21255767/)]
18. Furmark T, Carlbring P, Hedman E, Sonnenstein A, Clevberger P, Bohman B, et al. Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *Br J Psychiatry* 2009 Nov;195(5):440-447 [FREE Full text] [doi: [10.1192/bjp.bp.108.060996](https://doi.org/10.1192/bjp.bp.108.060996)] [Medline: [19880935](https://pubmed.ncbi.nlm.nih.gov/19880935/)]
19. 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](https://doi.org/10.1007/s12160-009-9133-4)] [Medline: [19802647](https://pubmed.ncbi.nlm.nih.gov/19802647/)]
20. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
21. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005;7(1):e7 [FREE Full text] [doi: [10.2196/jmir.7.1.e7](https://doi.org/10.2196/jmir.7.1.e7)] [Medline: [15829479](https://pubmed.ncbi.nlm.nih.gov/15829479/)]
22. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. *Br J Clin Psychol* 2010 Nov;49(Pt 4):455-471. [doi: [10.1348/014466509X472138](https://doi.org/10.1348/014466509X472138)] [Medline: [19799804](https://pubmed.ncbi.nlm.nih.gov/19799804/)]
23. Dasgupta A, Lawson KA, Wilson JP. Evaluating equivalence and noninferiority trials. *Am J Health Syst Pharm* 2010 Aug 15;67(16):1337-1343. [doi: [10.2146/ajhp090507](https://doi.org/10.2146/ajhp090507)] [Medline: [20689122](https://pubmed.ncbi.nlm.nih.gov/20689122/)]
24. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons, Inc; 1987.
25. Schafer JL. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall; 1997.
26. Spek V, Cuijpers P, Nyklíček I, Smits N, Riper H, Keyzer J, et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. *Psychol Med* 2008 May;38(5):635-639. [doi: [10.1017/S0033291707002590](https://doi.org/10.1017/S0033291707002590)] [Medline: [18205965](https://pubmed.ncbi.nlm.nih.gov/18205965/)]
27. Puhani P. The Heckman Correction for sample selection and its critique. *J Economic Surveys* 2000 Feb;14(1):53-68. [doi: [10.1111/1467-6419.00104](https://doi.org/10.1111/1467-6419.00104)]
28. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data*. New York: Springer; 2000.
29. Houck PR, Mazumdar S, Koru-Sengul T, Tang G, Mulsant BH, Pollock BG, et al. Estimating treatment effects from longitudinal clinical trial data with missing values: comparative analyses using different methods. *Psychiatry Res* 2004 Dec 15;129(2):209-215. [doi: [10.1016/j.psychres.2004.08.001](https://doi.org/10.1016/j.psychres.2004.08.001)] [Medline: [15590048](https://pubmed.ncbi.nlm.nih.gov/15590048/)]
30. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? *Clin Psychol Rev* 2011 Feb;31(1):89-103. [doi: [10.1016/j.cpr.2010.09.008](https://doi.org/10.1016/j.cpr.2010.09.008)] [Medline: [21130939](https://pubmed.ncbi.nlm.nih.gov/21130939/)]
31. Andersson G, Bergström J, Carlbring P, Lindfors N. The use of the Internet in the treatment of anxiety disorders. *Curr Opin Psychiatry* 2005 Jan;18(1):73-77. [Medline: [16639187](https://pubmed.ncbi.nlm.nih.gov/16639187/)]
32. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. *Clin Psychol Rev* 2009 Jul;29(5):410-420 [FREE Full text] [doi: [10.1016/j.cpr.2009.04.001](https://doi.org/10.1016/j.cpr.2009.04.001)] [Medline: [19403215](https://pubmed.ncbi.nlm.nih.gov/19403215/)]
33. Mental Health Online. Victoria, Australia: National eTherapy Centre, Swinburne University of Technology URL: <https://www.anxietyonline.org.au/default.aspx> [accessed 2014-09-28] [WebCite Cache ID 6SwJtAY5N]

34. Al-Asadi AM, Klein B, Meyer D. Pretreatment attrition and formal withdrawal during treatment and their predictors: an exploratory study of the anxiety online data. *J Med Internet Res* 2014;16(6):e152 [FREE Full text] [doi: [10.2196/jmir.2989](https://doi.org/10.2196/jmir.2989)] [Medline: [24938311](https://pubmed.ncbi.nlm.nih.gov/24938311/)]
35. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 2003 Feb;60(2):184-189. [Medline: [12578436](https://pubmed.ncbi.nlm.nih.gov/12578436/)]
36. Klein B, Meyer D, Austin DW, Kyrios M. Anxiety online: a virtual clinic: preliminary outcomes following completion of five fully automated treatment programs for anxiety disorders and symptoms. *J Med Internet Res* 2011;13(4):e89 [FREE Full text] [doi: [10.2196/jmir.1918](https://doi.org/10.2196/jmir.1918)] [Medline: [22057287](https://pubmed.ncbi.nlm.nih.gov/22057287/)]
37. Furukawa TA, Kessler RC, Slade T, Andrews G. The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychol Med* 2003 Feb;33(2):357-362. [Medline: [12622315](https://pubmed.ncbi.nlm.nih.gov/12622315/)]
38. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009 Nov;41(4):1149-1160. [doi: [10.3758/BRM.41.4.1149](https://doi.org/10.3758/BRM.41.4.1149)] [Medline: [19897823](https://pubmed.ncbi.nlm.nih.gov/19897823/)]
39. Heckman JJ. Sample selection bias as a specification error. *Econometrica* 1979;47(1):153-161. [doi: [10.3386/w0172](https://doi.org/10.3386/w0172)]
40. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L Erlbaum Associates; 1988.
41. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods* 1996 Jun;1(2):170-177.
42. Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. *Am J Psychiatry* 2006 Dec;163(12):2119-2125. [doi: [10.1176/appi.ajp.163.12.2119](https://doi.org/10.1176/appi.ajp.163.12.2119)] [Medline: [17151163](https://pubmed.ncbi.nlm.nih.gov/17151163/)]
43. Ruwaard J, Broeksteeg J, Schrieken B, Emmelkamp P, Lange A. Web-based therapist-assisted cognitive behavioral treatment of panic symptoms: a randomized controlled trial with a three-year follow-up. *J Anxiety Disord* 2010 May;24(4):387-396. [doi: [10.1016/j.janxdis.2010.01.010](https://doi.org/10.1016/j.janxdis.2010.01.010)] [Medline: [20227241](https://pubmed.ncbi.nlm.nih.gov/20227241/)]
44. Titov N, Andrews G, Choi I, Schwencke G, Johnston L. Randomized controlled trial of web-based treatment of social phobia without clinician guidance. *Aust NZ J Psychiatry* 2009 Jan;43(10):913-919. [doi: [10.1080/00048670903179160](https://doi.org/10.1080/00048670903179160)]
45. Macgill AR. Memo on parent and teen Internet use. Washington, DC: Pew Internet & American Life Project; 2007 Oct 24. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2007/PIP_Teen_Parents_data_memo_Oct2007.pdf.pdf [accessed 2014-09-28] [WebCite Cache ID 6SwKq4Bv4]
46. Johansson R, Sjöberg E, Sjögren M, Johnsson E, Carlbring P, Andersson T, et al. Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: a randomized controlled trial. *PLoS One* 2012;7(5):e36905 [FREE Full text] [doi: [10.1371/journal.pone.0036905](https://doi.org/10.1371/journal.pone.0036905)] [Medline: [22615841](https://pubmed.ncbi.nlm.nih.gov/22615841/)]
47. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. *Journal of Technology in Human Services* 2008 Jul 03;26(2-4):109-160. [doi: [10.1080/15228830802094429](https://doi.org/10.1080/15228830802094429)]
48. Titov N, Andrews G, Choi I, Schwencke G, Mahoney A. Shyness 3: randomized controlled trial of guided versus unguided Internet-based CBT for social phobia. *Aust N Z J Psychiatry* 2008 Dec;42(12):1030-1040. [doi: [10.1080/00048670802512107](https://doi.org/10.1080/00048670802512107)] [Medline: [19016091](https://pubmed.ncbi.nlm.nih.gov/19016091/)]
49. Brown C, Conner KO, Copeland VC, Grote N, Beach S, Battista D, et al. DEPRESSION STIGMA, RACE, AND TREATMENT SEEKING BEHAVIOR AND ATTITUDES. *J Community Psychol* 2010 Apr;38(3):350-368 [FREE Full text] [doi: [10.1002/jcop.20368](https://doi.org/10.1002/jcop.20368)] [Medline: [21274407](https://pubmed.ncbi.nlm.nih.gov/21274407/)]
50. Corrigan P. How stigma interferes with mental health care. *Am Psychol* 2004 Oct;59(7):614-625. [doi: [10.1037/0003-066X.59.7.614](https://doi.org/10.1037/0003-066X.59.7.614)] [Medline: [15491256](https://pubmed.ncbi.nlm.nih.gov/15491256/)]
51. Shechtman Z, Vogel D, Maman N. Seeking psychological help: a comparison of individual and group treatment. *Psychother Res* 2010 Jan;20(1):30-36. [doi: [10.1080/10503300903307648](https://doi.org/10.1080/10503300903307648)] [Medline: [19856235](https://pubmed.ncbi.nlm.nih.gov/19856235/)]
52. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA* 2004 Jun 2;291(21):2581-2590. [doi: [10.1001/jama.291.21.2581](https://doi.org/10.1001/jama.291.21.2581)] [Medline: [15173149](https://pubmed.ncbi.nlm.nih.gov/15173149/)]
53. Mojtabai R. National trends in mental health disability, 1997-2009. *Am J Public Health* 2011 Nov;101(11):2156-2163. [doi: [10.2105/AJPH.2011.300258](https://doi.org/10.2105/AJPH.2011.300258)] [Medline: [21940913](https://pubmed.ncbi.nlm.nih.gov/21940913/)]
54. Olfson M, Marcus SC. National trends in outpatient psychotherapy. *Am J Psychiatry* 2010 Dec;167(12):1456-1463. [doi: [10.1176/appi.ajp.2010.10040570](https://doi.org/10.1176/appi.ajp.2010.10040570)] [Medline: [20686187](https://pubmed.ncbi.nlm.nih.gov/20686187/)]
55. Caldwell TM, Jorm AF, Dear KB. Suicide and mental health in rural, remote and metropolitan areas in Australia. *Med J Aust* 2004 Oct 4;181(7 Suppl):S10-S14. [Medline: [15462636](https://pubmed.ncbi.nlm.nih.gov/15462636/)]
56. Nicholson LA. Rural mental health. *Advances in Psychiatric Treatment* 2008 Jul 01;14(4):302-311 [FREE Full text] [doi: [10.1192/apt.bp.107.005009](https://doi.org/10.1192/apt.bp.107.005009)]
57. Nguyen D. unpublished PhD thesis. Victoria: Swinburne University; 2013.

Abbreviations

CBT: cognitive behavioral therapy
DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders (4th Edition, Text Revision)
e-PASS: electronic psychological assessment screening system
GAD: generalized anxiety disorder
MDE: major depressive episode
OCD: obsessive-compulsive disorder
PD/A: panic disorder with or without agoraphobia
PTSD: posttraumatic stress disorder
SAD: social anxiety disorder

Edited by G Eysenbach; submitted 04.05.14; peer-reviewed by C Botella, C Donovan; comments to author 27.05.14; revised version received 22.06.14; accepted 16.08.14; published 14.10.14.

Please cite as:

AL-Asadi AM, Klein B, Meyer D

Posttreatment Attrition and Its Predictors, Attrition Bias, and Treatment Efficacy of the Anxiety Online Programs

J Med Internet Res 2014;16(10):e232

URL: <http://www.jmir.org/2014/10/e232/>

doi: [10.2196/jmir.3513](https://doi.org/10.2196/jmir.3513)

PMID: [25316533](https://pubmed.ncbi.nlm.nih.gov/25316533/)

©Ali M AL-Asadi, Britt Klein, Denny Meyer. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 14.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Understanding Usage of a Hybrid Website and Smartphone App for Weight Management: A Mixed-Methods Study

Leanne G Morrison¹, PhD; Charlie Hargood², PhD; Sharon Xiaowen Lin³, PhD; Laura Dennison¹, PhD; Judith Joseph¹, PhD; Stephanie Hughes¹, MSc; Danius T Michaelides², PhD; Derek Johnston⁴, PhD; Marie Johnston⁵, PhD; Susan Michie⁶, DPhil; Paul Little⁷, FMedSci; Peter WF Smith³, PhD; Mark J Weal², PhD; Lucy Yardley¹, PhD

¹Centre for Applications of Health Psychology, Academic Unit of Psychology, University of Southampton, Southampton, United Kingdom

²Electronics and Computer Science, University of Southampton, Southampton, United Kingdom

³Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, United Kingdom

⁴School of Psychology, University of Aberdeen, Aberdeen, United Kingdom

⁵Institute of Applied Health Sciences, College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, United Kingdom

⁶Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

⁷Primary Care and Population Sciences, University of Southampton, Southampton, United Kingdom

Corresponding Author:

Leanne G Morrison, PhD

Centre for Applications of Health Psychology

Academic Unit of Psychology

University of Southampton

Highfield Campus

Highfield

Southampton, SO17 1BJ

United Kingdom

Phone: 44 2380597222 ext 7222

Fax: 44 2380598721

Email: l.morrison@soton.ac.uk

Abstract

Background: Advancements in mobile phone technology offer huge potential for enhancing the timely delivery of health behavior change interventions. The development of smartphone-based health interventions (apps) is a rapidly growing field of research, yet there have been few longitudinal examinations of how people experience and use these apps within their day-to-day routines, particularly within the context of a hybrid Web- and app-based intervention.

Objective: This study used an in-depth mixed-methods design to examine individual variation in (1) impact on self-reported goal engagement (ie, motivation, self-efficacy, awareness, effort, achievement) of access to a weight management app (POWeR Tracker) when provided alongside a Web-based weight management intervention (POWeR) and (2) usage and views of POWeR Tracker.

Methods: Thirteen adults were provided access to POWeR and were monitored over a 4-week period. Access to POWeR Tracker was provided in 2 alternate weeks (ie, weeks 1 and 3 or weeks 2 and 4). Participants' goal engagement was measured daily via self-report. Mixed effects models were used to examine change in goal engagement between the weeks when POWeR Tracker was and was not available and whether the extent of change in goal engagement varied between individual participants. Usage of POWeR and POWeR Tracker was automatically recorded for each participant. Telephone interviews were conducted and analyzed using inductive thematic analysis to further explore participants' experiences using POWeR and POWeR Tracker.

Results: Access to POWeR Tracker was associated with a significant increase in participants' awareness of their eating ($\beta_1=0.31$, $P=.04$) and physical activity goals ($\beta_1=0.28$, $P=.03$). The level of increase varied between individual participants. Usage data showed that participants used the POWeR website for similar amounts of time during the weeks when POWeR Tracker was (mean 29 minutes, SD 31 minutes) and was not available (mean 27 minutes, SD 33 minutes). POWeR Tracker was mostly accessed in short bursts (mean 3 minutes, SD 2 minutes) during convenient moments or moments when participants deemed the intervention content most relevant. The qualitative data indicated that nearly all participants agreed that it was more convenient to access

information on-the-go via their mobiles compared to a computer. However, participants varied in their views and usage of the Web- versus app-based components and the informational versus tracking tools provided by POWeR Tracker.

Conclusions: This study provides evidence that smartphones have the potential to improve individuals' engagement with their health-related goals when used as a supplement to an existing online intervention. The perceived convenience of mobile access to information does not appear to deter use of Web-based interventions or strengthen the impact of app access on goal engagement. A mixed-methods design enabled exploration of individual variation in daily usage of the app-based tools.

(*J Med Internet Res* 2014;16(10):e201) doi:[10.2196/jmir.3579](https://doi.org/10.2196/jmir.3579)

KEYWORDS

qualitative research; weight loss; behavioral research; mobile apps; Internet; health; program acceptability; behavior; mixed-methods

Introduction

Background

Over the past decade, there has been a proliferation of digital technologies to deliver interventions designed to support health behavior change, including computer- and Internet-based platforms [1-4], social media and online social networks [5,6], and mobile phones and other handheld devices [7,8]. Feature-rich smartphones arguably offer unique advantages over these forms of digital delivery given their apparent ubiquity and widespread penetration within individuals' daily lives combined with the opportunity to harness their context-aware sensing capabilities [9-11]. Smartphone apps have been used in a number of different ways to promote and support health, including (but not limited to) automated prompts and reminders, information provision, self-monitoring and tracking, remote monitoring by health professionals, and incorporation of social networks [9].

There has been a surge in the development of apps to support a diverse range of health issues, including health promotion and disease prevention (eg, nutrition and physical activity [12-16], weight management [17,18], protective sun behaviors [19], substance use [20,21]), self-management of chronic physical conditions (eg, diabetes [22-24], pain [25], asthma [26]), and self-management of mental health [27] (eg, anxiety, stress, and well-being [28-34], depression [35-39], schizophrenia [40]). Despite the proliferation of apps, research on the feasibility and effectiveness of app-based health interventions is still at a relatively early stage and, to date, has largely focused on exploring user needs, the development of content specific to a given health behavior, and/or identifying specific usability issues.

Apps can and have been designed as stand-alone health interventions. An alternative hybrid model is to provide larger packages of Web-based information and advice supplemented by on-the-go mobile-based access to specific intervention components [14,18,21,28]. There has been comparatively little in-depth examination of how individuals view, use, and react to a hybrid intervention model. For example, do individuals differ in their preference for using different digital formats (ie, Web vs app)? [41] Such insights are vital for understanding in what contexts and for whom specific forms of digital delivery may be more successful at promoting optimal usage of and adherence to health behavior interventions. For example, apps have shown particular promise in enabling prompting [15] and

self-monitoring of health-related behavior [17,18,42-44] as well as increasing users' awareness of health-related goals and behaviors [15,16,32,45,46].

There also appears to be little adequate exploration of how individuals actually use their mobile phones and integrate health apps in their day-to-day routines using longitudinal case study-based approaches [47,48]. The apparent ubiquity of mobile phones in our everyday lives suggests that apps will improve the accessibility, reach, and convenience of health interventions resulting in greater uptake, adherence, and subsequent health improvement. Mobile and app-based interventions appear to be welcomed by individuals living with chronic health conditions [22,24,49,50]. For example, diabetic adolescents suggested that a mobile app could be integrated within their daily routine to facilitate blood glucose monitoring and provide medication and appointment reminders [49]. However, apps may not always be more convenient and better suited to the delivery of self-monitoring components than Web-based interventions.

Qualitative studies have shown that while people are receptive to the possibility of using mobile health apps, these apps may be easily discarded [51], particularly if they are not designed for flexible, quick, and effort-free use [22,43,52,53], and that there are certain contexts in which app use can be perceived as inappropriate or embarrassing (eg, during working hours or eating out with peers [16,22]). There is also evidence to suggest that app-based tracking of behavior is not easily incorporated into individuals' daily routines if unprompted [16], yet automated mobile prompts appear to be disliked by users when they are received at inappropriate moments [42] or too frequently [19]. These insights have largely been derived from participants' hypothetical perceptions of health apps, experience with commercial apps, or brief interactions with app-based health behavior change interventions. Longitudinal research of actual use can improve our understanding of when, why, and how individuals experience and make use of health apps, or indeed their mobile phones, within their daily lives. This is needed for designing optimal user interfaces that enable quick and easy access to the core app components used on a day-to-day basis [54] as well as to inform the tailoring of automated prompts to ensure they are received at the right time with the right content.

We are aware of few studies that have combined quantitative analyses of intervention usage and change in self-reported outcomes with in-depth qualitative research to examine and compare individual differences in the impact of and experience

of health apps. This study used a novel mixed-methods design to examine the impact of, usage, and views of a weight management app (POWeR Tracker) that was provided alongside a Web-based weight management intervention (POWeR). Facilitating users to set personal healthy eating and physical activity goals is a key behavior change technique (BCT) incorporated in POWeR. The Health Action Process Approach argues that behavior change (in this case pursuit of weight management goals) occurs through 2 phases: motivational and volitional [55]. Coping self-efficacy (the extent to which an individual feels confident in overcoming barriers to goal pursuit) and action control (awareness of intended behaviors and self-regulatory effort) are identified as key predictors of behavior [56]. This study first examined whether providing access to POWeR Tracker enhanced participants' self-reported goal engagement and, if so, whether/how the extent of enhancement varied between participants. Goal engagement was conceptualized in this study as motivation for goal pursuit, coping self-efficacy, action control (awareness of goals, effort toward goals), and achievement of goals. Variation in each participant's views of POWeR Tracker and their usage of it on a day-to-day basis were then examined to further understand how and why it was used and to offer explanations for any individual variation observed in the impact of POWeR Tracker access on levels of goal engagement. The efficacy of the Web-based POWeR intervention and participants' views of the specific content provided by the Web-based intervention have been the focus of previous qualitative studies [57] and randomized controlled trials within health care [58] and community settings [59].

The Interventions

POWeR

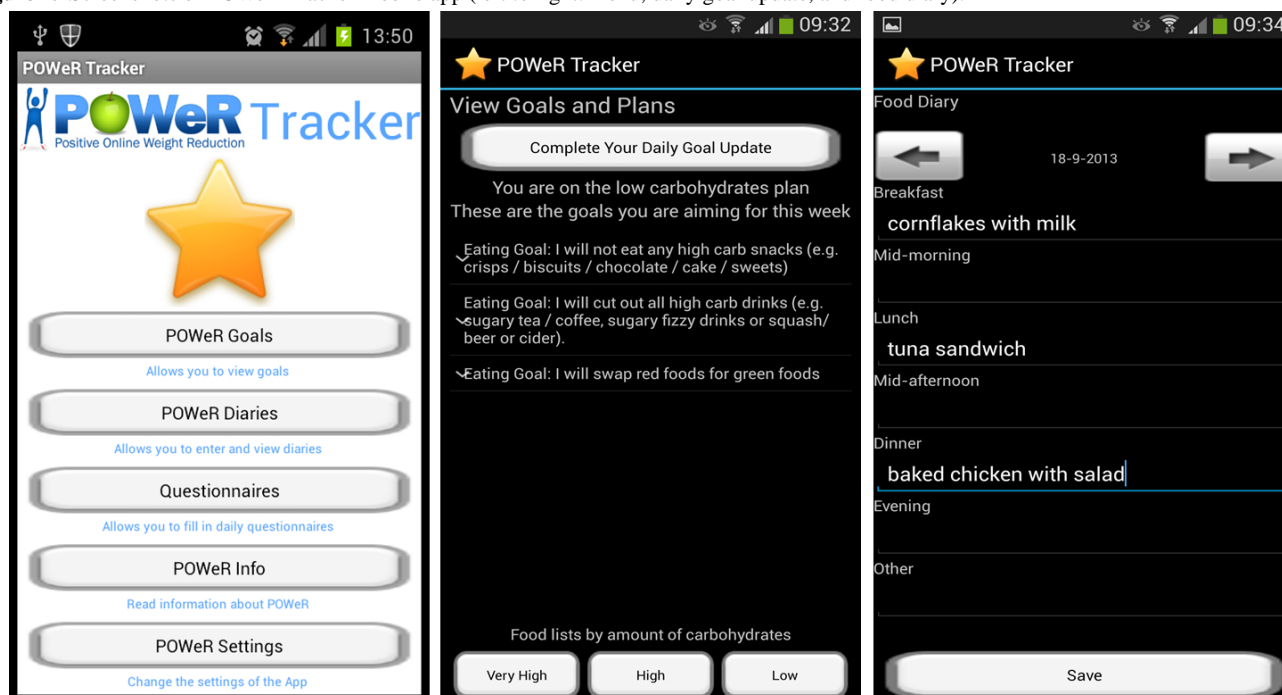
Positive Online Weight Reduction (POWeR) is a Web-based weight management intervention that was developed using the LifeGuide authoring software [60]. POWeR offers a flexible, nonprescriptive approach to weight management to foster autonomy and support users to adopt healthy behaviors that will empower them to maintain long-term weight management. POWeR is delivered over 12 sessions that become available to users on a weekly basis. Each session is comprised of a range of "POWeR tools" that support the development of self-regulatory skills, "POWeR stories" that model successful weight management strategies, and links to further sources of information and advice. A detailed overview of the iterative

development and qualitative piloting of the Web-based POWeR intervention is provided elsewhere [57].

The first 3 sessions are designated as core sessions. Session 1 introduces the POWeR approach, guides users to select goals consistent with either a low-calorie or a low-carbohydrate eating plan, and review their personal motivations for losing weight. Session 2 provides tips on getting support for weight management. Session 3 guides users to select goals consistent with either a walking or mixed physical activity plan. Setting personalized goals (and developing detailed plans of how to meet those goals), reviewing goal progress, and getting Web-based feedback on goal achievement is an essential element of the POWeR intervention. From session 2 onward, participants are required to log their weight and review their goals before accessing new session content. During sessions 4-11, users are invited to choose topics of interest (eg, controlling cravings, dealing with slip-ups, eating out) after completing their weight and goal review. The final session of POWeR provides information and advice on how users can maintain their weight management over the longer term.

POWeR Tracker

POWeR Tracker is an Android mobile phone app that offers users of POWeR the opportunity to keep track of their personal POWeR goals via their mobile phone (refer to [Multimedia Appendix 1](#) for a detailed overview). POWeR Tracker is supplemental to, rather than a duplication of, the Web-based POWeR intervention. It provides a range of tools that are designed to enhance users' awareness of and motivation to work toward their personal POWeR goals between the Web-based weekly POWeR sessions (see [Figure 1](#)). POWeR Tracker provides 2 types of tools: informational and self-monitoring. Informational tools include options to view one's goals and plans and access selected content that was introduced during the first Web-based POWeR session, such as food lists of items that are high or low in calories or carbohydrates, one's personal reasons to lose weight card, and advice on how keeping diaries can support weight management. Self-monitoring tools include options to receive personalized feedback on progress toward goals via completion of a daily goal update and to complete food and physical activity diaries. POWeR Tracker also offers users the ability to set up automated notifications that provide reminders to view goals or complete a daily goal update at a time of their choosing. Participants are free to cancel notifications at any time.

Figure 1. Screenshots of POWeR Tracker mobile app (left to right: menu, daily goal update, and food diary).

Methods

Recruitment

A volunteer sample of 13 participants was recruited using paper-based advertisements placed around the campus of the University of Southampton, UK. Eligible participants were required to have a body mass index (BMI) of at least 23, own an Android mobile phone, and have no pre-existing health conditions that would impede modification of nutrition or physical activity. Recruitment ceased when no substantially different insights of participants' experiences of POWeR Tracker were derived (ie, when saturation was achieved).

Design and Procedure

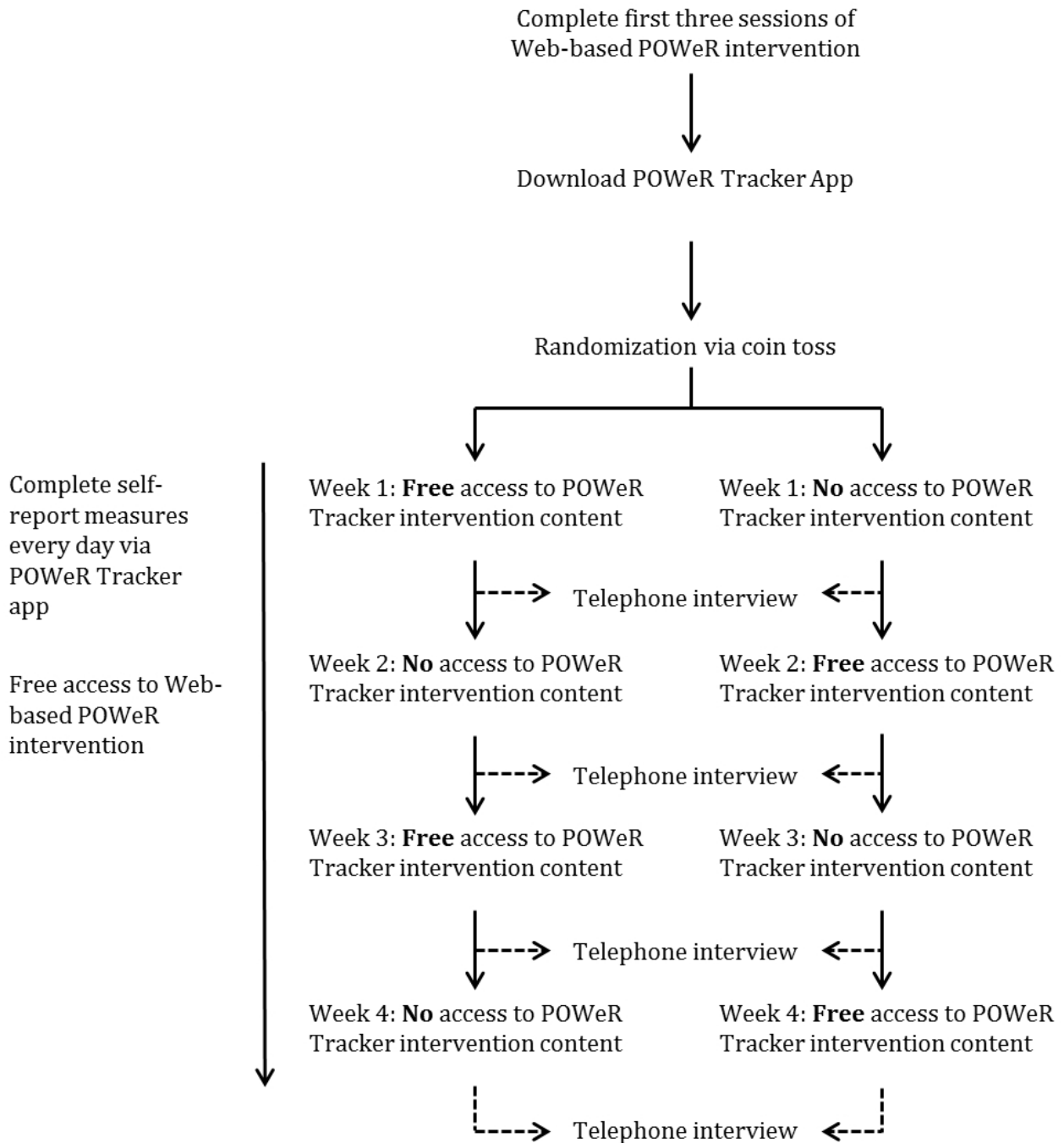
Data collection took place between August 2012 and August 2013 and was approved by the University of Southampton ethics committee and research governance office. Each participant was required to select their personal eating and physical activity plans and goals during the first 3 Web-based sessions of the POWeR intervention. After completing the first 3 Web-based sessions, each participant was invited to download the POWeR Tracker app and was then monitored over a 4-week period (see [Figure 2](#)). During this period, participants could continue to freely use the Web-based POWeR intervention.

When referring to the POWeR Tracker app, the phrase "intervention content" refers to the information, advice, and tools provided, whereas the phrase "daily questionnaires" refers to the self-report study measures. Access to the intervention content provided by the POWeR Tracker app was restricted and alternated on a weekly basis for each participant. Participants could either access the intervention content in weeks 1 and 3 or in weeks 2 and 4. The order of first access to the POWeR

Tracker intervention content was counterbalanced across participants; participants were randomized via coin toss in blocks of 4 to receive first access during the first or second week of the study.

Participants were also required to complete a number of self-report measures of goal engagement (daily questionnaires) every day for all 4 weeks of the study via the POWeR Tracker app. In-line with the Health Action Process Approach [56], goal engagement was conceptualized in this study as motivation for goal pursuit, coping self-efficacy, action control (awareness, effort), and achievement of goals. To prevent backfilling, daily questionnaires could only be completed between 5 pm that day and 11 am the following morning. Semistructured telephone interviews were conducted at the end of each week to discuss each participant's experiences of using POWeR and POWeR Tracker (see [Multimedia Appendix 2](#) for the interview schedule). The interview schedule was initially developed collaboratively by LM, LD, and LY, but evolved over the course of the study based on responses from participants. Each interview was conducted by LM, LD, JJ, or SH and lasted between 5 and 38 minutes (mean 15 minutes). All study procedures and materials were initially pilot-tested with 1 user who was included in subsequent analyses (P1a). Participants were reimbursed for their time with either cash (£75) or research participation credits. Reimbursement was conditional on completion of the daily self-report measures and participation in the weekly telephone interviews. Reimbursement was not conditional on usage of the app- or Web-based intervention content. Participants were not provided with any training on how to register, download, or use POWeR/POWeR Tracker. However, participants were free to raise and receive advice for resolving any technical problems during the weekly telephone interviews.

Figure 2. Flowchart of study design.



Measures

Usage of POWeR and POWeR Tracker was recorded automatically for each participant using the LifeGuide software, including when, how long, and in what order particular pages or screens were viewed. All participants were informed that their usage of POWeR and POWeR Tracker would be recorded. During the first 3 core sessions of the Web-based POWeR intervention, users were guided to create 3 eating goals and 3 physical activity goals. Self-reported goal engagement was recorded every day during the 4 week study period via the POWeR Tracker app to assess whether access to the POWeR Tracker app enhanced (1) motivation for goal pursuit (goal motivation), (2) coping self-efficacy (goal self-efficacy) (3)

action control (goal awareness, goal effort), and (4) achievement of goals (goal achievement). Motivation, self-efficacy, awareness, and achievement were each measured for eating goals and physical activity goals using 3-item scales developed for this study (Cronbach alpha=.95-.99) (see [Multimedia Appendix 3](#)). Goal effort toward each of the 6 individual goals was measured using a 3-item scale developed by Louro et al [61] (Cronbach alpha=.87). All measures used a 7-point Likert scale that ranged from strongly disagree to strongly agree.

Analysis

Between-Week Differences in Goal Engagement

Statistical analyses were performed using the R software environment version 3.0.2 [62]. Two mixed effects models were

fitted to the combined daily self-report data from 12 participants to test whether there was a difference in the summary scores for each self-report measure (outcome variable) between the weeks when the POWeR Tracker app was and was not available (predictor variable) (Model 1) and, if so, whether the extent of those differences varied between individual participants (Model 2). Summary scores for each self-report measure were calculated by averaging the responses provided to each of the 3 items. One participant (P7a) was excluded from these analyses because he was unable to provide daily self-report data due to a technical fault with the POWeR Tracker app. Model 1 represented the experimental hypothesis specifying fixed (β_0) and random effects (σ_{ou}) for the intercept and a fixed (β_1) intervention effect, where β_0 denotes the average baseline value on day 1. Model 2 also allows for fixed (β_0) and random effects (σ_{ou}) for the intercept and a fixed (β_1) intervention effect, but with an additional random effects (σ_{iu}) on the intervention to allow for individual intervention effects. Model 2 was only fitted to data from the self-report measures that showed a significant intervention effect in Model 1. The fit of Model 1 versus Model 2 to the data was compared for self-report measures that showed a significant intervention effect. Likelihood ratio tests were used to select the preferred model fit for each self-report measure (ie, Model 1 vs Model 2). Model parameters were calculated for each individual participant where significant random effects of the intervention were observed. Both Model 1 and Model 2 allowed the individual level residuals to be correlated using an autoregressive process of order 1.

Website and App Usage Patterns

Usage of POWeR and POWeR Tracker was summarized for each participant to compare duration, frequency, and time of access to intervention content. Averages were also computed based on the combined usage data of all 13 participants to summarize how often and when the sample as a whole used POWeR and POWeR Tracker during the study period. Bivariate correlational analyses (Pearson r) were used where appropriate to examine whether observed patterns in participants' usage of POWeR and POWeR Tracker were statistically significant.

Telephone interviews were audio recorded and transcribed verbatim. Transcripts for each of the 13 participants were analyzed using inductive thematic analysis [63]. In the first phase of analysis, each participant was considered individually so as to remain sensitive to the nuances of each individual's experience of POWeR. This involved the development of initial codes to label each segment of text, which were then used to produce a set of themes that summarized the experiences for each participant. Analyzing each participant's data separately ensured that unique or "deviant" views and/or experiences of POWeR were preserved when seeking to synthesize and compare data across the whole sample (in phase 2). This

facilitated interpretation of individual variation observed for the impact of app access on goal engagement.

In the second phase, a thematic analysis of the themes generated for each participant (in phase 1) was conducted to generate a set of themes that captured the experiences of all 13 participants. This involved comparing the content of themes initially generated for each participant to merge and synthesize across participants while also preserving any differences between individuals in how they viewed and experienced POWeR. Participant identification numbers (eg, P1a, P11b) were used to protect the anonymity of participants, where "a" indicates first app access in week 1 and "b" indicates first app access in week 2.

Results

Overview

Presentation of the results will be provided in 4 sections. The first will describe the participant characteristics. The second will report on the significant associations between provision of the POWeR Tracker app and change in self-reported goal engagement. The final 2 sections will outline participants' usage and views of the POWeR Tracker app and the concurrent Web- and app-based delivery of POWeR.

Participants

A total of 13 (6 male and 7 female) healthy adults aged 18-52 years (median 27 years) participated. The BMI of participants ranged from 23.69 to 38.51 kg/m² (median 26 kg/m²). All participants either completed or were currently studying for a degree. On average, participants reported using their mobile phone from between 1-16 hours per day (median 2 hours). Most of the participants cited more than 1 motivation for signing up for the study. The most common motivations were to lose weight and get fitter (both cited by 8/13, 62%), and to learn about a healthy lifestyle, earn money, and contribute to current research (all cited by 3/13, 23%). Only 1 participant (8%) reported that he/she was interested in trying new apps.

Impact of POWeR Tracker App on Goal Engagement

Access to the POWeR Tracker app was associated with a significant increase in self-reported motivation, self-efficacy, awareness, and achievement of eating goals and a significant increase in self-reported awareness of physical activity goals. There were no significant differences in self-reported goal effort (for eating or physical goals) or in self-reported motivation, self-efficacy, and achievement of physical activity goals between the weeks when the POWeR Tracker app was and was not available. Table 1 presents the estimates for Model 1 along with the standard error for the estimate of β_1 and the P value for the Wald test of the null hypothesis that $\beta_1=0$.

Table 1. Estimates for Model 1 testing for fixed intervention effect.

Measure	β_0^a	β_1				σ_{ou}^c	ρ^d	$R^2, \%^e$	
		β_1^b	SE	P			Marginal	Conditional	
Eating goals									
Motivation	4.68	0.42	0.15	.01	1.41	.05	1.30	60	
Self-efficacy	4.53	0.34	0.15	.03	1.47	.11	.83	64	
Awareness	5.26	0.31	0.15	.04	1.19	.22	.96	58	
Effort, goal 1 ^f	4.98	0.22	0.17	.18	1.29	.02	.38	52	
Effort, goal 2 ^f	5.07	0.24	0.16	.15	1.27	-.04	.43	49	
Effort, goal 3 ^f	4.84	0.10	0.24	.69	0.88	.40	.07	24	
Achievement	4.64	0.32	0.17	.06	1.38	.21	.79	60	
Physical activity goals									
Motivation	4.86	0.12	0.18	.52	1.40	.34	.10	58	
Self-efficacy	4.85	-0.09	0.18	.64	1.35	.16	.05	52	
Awareness	5.14	0.28	0.13	.03	1.38	.06	.71	68	
Effort, goal 1 ^f	4.80	-0.02	0.30	.95	1.02	.19	.00	20	
Effort, goal 2 ^f	4.66	-0.20	0.26	.45	1.42	.16	.18	38	
Effort, goal 3 ^f	4.31	0.13	0.30	.67	1.42	.43	.07	35	
Achievement	4.56	0.14	0.21	.52	1.39	.32	.12	51	

^a β_0 denotes average baseline value on day 1, where minimum possible score is 1 (strongly disagree) and maximum possible score is 7 (strongly agree).

^b β_1 denotes the average change in scores for all participants between the weeks when the POWeR Tracker app was and was not available.

^c σ_{ou} denotes standard deviation of random effects for changes in average baseline value on day 1.

^dFor autocorrelation.

^e R^2 marginal denotes the proportion of total variation in each measure explained by access to POWeR Tracker; R^2 conditional denotes the proportion of total variation in each measure explained by access to POWeR Tracker and individual variability in self-report responses.

^fEffort, goal 1-3 denotes each eating and physical activity goal set by participants.

The magnitude of change in self-reported awareness of eating and physical activity goals and self-reported achievement of eating goals as a result of access to the POWeR Tracker app varied significantly between participants. However, no significant variations between participants were observed for change in self-reported motivation and self-efficacy for eating goals. Table 2 presents the estimates for Model 2 along with

standard error (SE) of β_1 , the *P* value for the Wald test of the null hypothesis that $\beta_1=0$, the standard deviation of both the random intercept (σ_{0u}) and random effects for the intervention (σ_{1u}), and the correlation (ρ_{01u}) between the 2 random effects. Likelihood ratio (LR) test statistics for comparing Model 1 and Model 2 and their respective *P* values are also displayed in Table 2.

Table 2. Estimates for Model 2 testing for individual intervention effects.

Measure	β_0^a	β_1			σ_{ou}^c	σ_{1u}	ρ_{01u}	ρ	Likelihood ratio	
		β_1^b	SE	P					LR	P
Eating goals										
Motivation	4.61	0.48	0.24	.05	1.59	0.63	-.55	.01	3.97	.14
Self-efficacy	4.52	0.35	0.18	.05	1.49	0.32	-.16	.09	0.38	.83
Awareness	5.13	0.45	0.35	.19	1.63	1.07	-.76	.06	20.59	<.001
Achievement	4.55	0.42	0.32	.19	1.57	0.96	-.47	.07	10.52	.01
Physical activity goals										
Awareness	5.11	0.31	0.22	.16	1.61	0.62	-.70	-.01	8.76	.01

^a β_0 denotes average baseline value on day 1, where minimum possible score is 1 (strongly disagree) and maximum possible score is 7 (strongly agree).

^b β_1 denotes change in scores between the weeks when the POWeR Tracker app was and was not available where the magnitude of change may be different for each participant.

^c σ_{ou} denotes standard deviation from average baseline value on day 1.

^d σ_{1u} denotes standard deviation of random effects for changes in scores between the weeks when the POWeR Tracker app was and was not available.

Participants P1a, P8b, and P11b reported the greatest improvement in awareness and achievement of goals during the weeks when access to POWeR Tracker was provided. In contrast, participants P3a, P5a, P6a, and P13b reported little or

no improvement in their awareness and achievement of goals. Table 3 presents the estimated intercept (β_{0i}) and intervention effects (β_{1i}) for each individual participant where there was a significant random effect for the intervention.

Table 3. Estimated individual intercept (β_{0i})^a and intervention effects (β_{1i})^b for each participant.

Participant	Goal awareness (eating)		Goal achievement (eating)		Goal awareness (physical activity)	
	β_{0i}	β_{1i}	β_{0i}	β_{1i}	β_{0i}	β_{1i}
First app access: week 1 of 4						
P1a	1.79	3.32	2.23	2.67	2.35	1.66
P2a	3.57	0.48	3.65	0.28	3.27	0.52
P3a	3.83	0.04	1.79	-0.38	4.07	0.02
P4a	5.93	0.34	5.43	0.54	6.11	0.15
P5a	6.05	0.04	5.11	-0.13	6.89	-0.20
P6a	6.87	-0.03	4.98	-0.14	6.99	-0.05
First app access: week 2 of 4						
P8b	3.86	0.78	3.18	1.11	4.05	0.84
P9b	5.84	-0.13	5.60	0.19	3.54	0.18
P10b	4.66	0.28	5.27	-0.06	5.41	0.16
P11b	6.27	0.49	4.74	0.76	6.14	0.24
P12b	6.21	0.14	6.04	0.25	5.78	0.33
P13b	6.70	-0.32	6.57	0.03	6.71	-0.12

^a β_0 denotes baseline value on day 1, where minimum possible score is 1 (strongly disagree) and maximum possible score is 7 (strongly agree).

^b β_1 denotes change in scores between the weeks when the POWeR Tracker app was and was not available.

Usage of POWeR Tracker App

On average, participants spent 29 minutes (SD 21 minutes) using the app-based intervention content during the 4 week study period. As a group, participants only accessed the app-based intervention content on an average of 9 of 14 (64%)

days it was available. Participants also spent a longer total time using the app-based intervention content during the first week of access (mean 18 minutes, SD 11 minutes) compared to the second week of access (mean 12 minutes, SD 14 minutes).

Several individual participants did, however, access the app on most of the days it was available (see Table 4). These participants tended to request more notifications and made greater use of the app-based tracking tools, particularly the daily goal update, as compared to those participants who accessed the app on a fewer number of days (see Figure 3). A significant positive correlation was observed between total duration of app use and effect of app access on awareness of eating goals

($r=.585, P=.046$), achievement of eating goals ($r=.620, P=.03$), but not awareness of physical activity goals ($r=.498, P=.10$). However, more regular use the tracking-based tools was not significantly related to how strongly app access influenced goal engagement (awareness of eating goals: $r=.525, P=.08$; achievement of eating goals: $r=.530, P=.08$; awareness of physical activity goals: $r=.387, P=.21$).

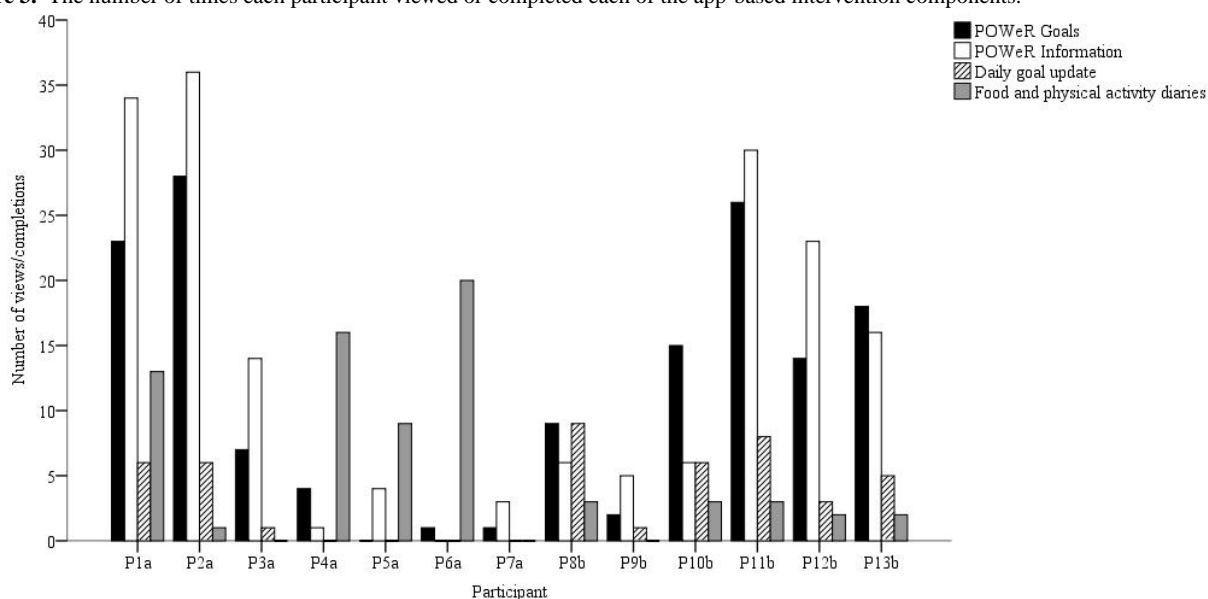
Table 4. Summary of participants' usage of the intervention content provided by the POWeR Tracker app.^a

Participant	Duration of app use (minutes)				Days app used (0-14 days), n (%)	Notifications requested by participants, n		
	Total	First week	Second week	Per day, ^b mean		Total	View goals	Daily goal update
First app access: week 1 of 4								
P1a	64.30	20.38	43.92	5.36	12 (86)	4	2	2
P2a	53.57	37.82	15.75	5.55	11 (79)	1	1	0
P3a	13.20	9.40	3.80	2.20	6 (43)	0	0	0
P4a	47.27	35.3	11.97	6.75	7 (50)	0	0	0
P5a	15.48	9.53	5.95	2.21	7 (50)	0	0	0
P6a	29.33	25.62	3.72	2.44	12 (86)	0	0	0
P7a	1.32	1.32	0	0.33	4 (29)	0	0	0
First app access: week 2 of 4								
P8b	24.08	17.23	6.85	2.41	10 (71)	17	8	9
P9b	2.87	1.10	1.77	0.91	4 (29)	8	3	5
P10b	20.15	17.32	2.83	3.36	6 (43)	0	0	0
P11b	57.27	26.26	31.02	4.09	14 (100)	21	13	8
P12b	17.08	14.2	28.95	1.42	12 (86)	26	14	12
P13b	41.17	12.38	28.78	3.74	11 (79)	16	9	7

^aDue to technical issues with the app, notifications were only intermittently received by 5 participants (P1a, P2a, P3a, P9b, and P11b). Therefore, the reported number of notifications requested may underestimate these participants' usage of this component.

^bRepresents average duration of use only on days when the app-based intervention content was accessed.

Figure 3. The number of times each participant viewed or completed each of the app-based intervention components.



Perceived Advantages of App Access

Nearly all the participants (10/13, 77%) stated that it was more

convenient to access the app than the website because their mobile phones were always with them and could be accessed on-the-go at any time (see [Table 5](#)).

Table 5. Overview of themes identified from participant interviews.

Theme	Overview
Convenience and accessibility: short bursts of on-the-go access	The app was considered a convenient and accessible means of accessing content because a phone is portable and can be used on-the-go in any location. Participants were happy to use the website and app in tandem if they were perceived to provide different value in different contexts.
Constant reminder and repetition	The app provided a constant reminder of participants' goals and plans. This helped to maintain awareness of goals and keep them in mind. App-based reminders were typically not considered necessary on a daily or longer-term basis.
Motivational benefits of tracking	Participants reported motivational benefits from logging and tracking thoughts and behaviors via diaries, goal updates, and daily questionnaires. Use of tracking tools facilitated recognition of both progress and areas for future improvement.
Time-relevant use guided by lifestyle and routine	Usage of the app was tied to personal lifestyle and fluctuations in daily routine. Participants typically reported using the app during free moments or specific times when the app-based intervention content was most relevant, such as mealtimes.

The perceived convenience of the app appeared to resolve barriers to accessing Web-based intervention content; the app could be used during a spare 5 minutes while a chunk of allocated time was needed to complete a Web-based session. For example, P3a said:

Because it's kind of in front of me all the time and in my hand, it's easier for me if I am thinking about those kinds of issues, to look on the app rather than log on to a computer and go on the website.

An exception was P8b who reported that app-based access to intervention content was actually less convenient:

Actually I thought the website was more, for me anyway, was more accessible than the app...But that might be because in terms of my phone I kind of use my phone just for texting and calling but I've got my iPod for all the apps so I don't normally connect my phone with an app...I would actually physically have to remember to use it on my phone.

Views on the convenience and accessibility of POWeR Tracker did not, however, appear to be strongly tied to observed usage or the effect of the app on goal engagement. For example, P8b used the app comparatively more and showed one of the strongest intervention effects despite perceiving app access as less convenient than Web access. P3a used the app comparatively less and showed one of the weakest intervention effects despite perceiving app access as more convenient than Web access (see [Table 3](#)).

Most participants agreed that another primary benefit of the app was that it provided a constant reminder of one's POWeR goals (see [Table 5](#)). [Figure 3](#) illustrates that the static information screens (ie, goal lists, food lists) were the most frequently viewed parts of the app for more than half the participants (9/13, 69%). This constant reminder of goals was reported to improve focus, awareness, and motivation. For example, P1a said:

You are always going to have goals in the back of your mind and reasons why you want to do things, but having it written down somewhere where you can

just go and have a look at it is quite good. I mean that is quite a good thing to have because obviously if you write it down on a piece of paper you are most likely not to look at it well I don't think I would to have an app to hand 24/7 is quite good.

Participants appeared to differ by how frequently they reported needing these reminders. Some participants stated that reminders of goals or food lists were not needed on a daily or continued basis particularly once the relevant information had been committed to memory, as P2a explains:

But once I've got the list of foods, once I know my targets and my goals, there is really no reason to log in...once it's in my head I just don't feel the need to look this stuff up.

Participants discussed how regular use of the tracking tools, including the daily questionnaires, facilitated critical self-reflection on progress and prompted further goal-directed behavior. P11b explained how answering questions about their goals each night on a Likert scale provided a useful gauge for realizing when it was time to set a new goal:

When you [ve made]...your goal a habit and then you change it [for a new goal], I think that's the most profound thing [that's happened to me during the study], [POWeR] gives you suggestions...[that] you should try and cut out drinking sugary tea or eating chocolate and then you try and do it and then you find, oh I can actually [did] that. And after [a goal] become[s] a habit you...try and do something else.

Participants suggested that the appeal and experience of using the app-based tracking tools could be optimized by providing effort-free data entry for diaries (eg, drop-down menus, barcode scanners) and personalized feedback. For example, P4a commented that the app could provide more feedback in-line with the weight graph provided during the weekly Web-based review:

Maybe a bit more feedback from what you said over the week or something... because it takes in all this

information and then it could spit something back out.

Usage of POWeR Website

Participants continued to use the POWeR website despite the perception that app-based access was more convenient (see Table 6). Total duration of app use was significantly correlated with total duration of website use ($r=.77, P=.002$). On average, participants spent approximately 56 minutes (SD 44 minutes) using the POWeR website and completed 3 of 4 available

Web-based sessions during the 4-week study period. Although similar amounts of total time were spent on the website during the 2 weeks when participants did (mean 29 minutes, SD 31 minutes) and did not (mean 27 minutes, SD 33 minutes) have access to the app-based intervention content, time spent on the website during nonapp weeks was significantly correlated with effect of app access on awareness of eating goals ($r=.930, P<.001$), achievement of eating goals ($r=.849, P<.001$), and awareness of physical activity goals ($r=.867, P<.001$).

Table 6. Summary of participants' usage of the POWeR website during weeks when the app-based intervention content was and was not available.

Participant	Duration (minutes)			Session completion, n					
	All weeks	App weeks	Nonapp weeks	Weight and goal review completion			Extra topics viewed		
				All weeks	App weeks	Nonapp weeks	All weeks	App weeks	Nonapp weeks
First app access: week 1 of 4									
P1a	127.54	6.06	123.15	3	0	3	2	0	2
P2a	154.11	111.10	43.01	2	1	1	2	1	1
P3a	54.78	15.42	39.35	4	1	3	3	1	2
P4a	46.98	10.52	36.46	4	1	3	1	0	1
P5a	17.53	6.84	10.69	2	2	0	0	0	0
P6a	51.71	30.99	20.72	4	2	2	3	2	1
P7a	9.74	3.42	6.31	2	1	1	0	0	0
First app access: week 2 of 4									
P8b	67.06	25.32	41.74	3	1	2	3	1	2
P9b	26.24	22.16	4.08	2	1	1	1	1	0
P10b	0	0	0	0	0	0	0	0	0
P11b	68.96	44.49	24.48	3	2	1	3	2	1
P12b	40.60	33.21	7.39	3	2	1	3	2	1
P13b	67.14	67.14	0	2	2	0	1	1	0

Participants P1a, P2a, and P11b expressed the most positive perceptions of the website compared to the other participants. All 3 stated that they were happy to use both the website and app to access different intervention content. Their usage patterns reflect this—they all spent an above average amount of time using both the website and app. P2a and P11b also discussed how app- versus Web-based access offered value within different contexts. For example, P11b described using the app on-the-go for quick updates, but using the website for more intensive support when he felt that progress toward his goals was waning. P2a said:

If we were comparing the website and the app, I spent a lot of time across the two...that was all quite straightforward that worked well. Anything like recording the foods, the diaries, and the [physical] activities...I do prefer doing that on my computer and I've got time to do it. The app is...most useful [for] look[ing] up the lists of food and stuff like that.

In contrast, P9b and P10b commented that they would prefer to access the entirety of the POWeR intervention via the app.

For example, P10b stated that they did not need to look at the additional content provided by the website as they were making good progress toward their POWeR goals by using the app alone. P9b discussed the value of accessing “quick little snapshots of information” via the app in comparison to the long, text-heavy, Web-based sessions:

Apps are just a lot more instant and I tend to try and do it when I'm on-the-go if I can as well. When I'm out and about, [I] try and do bits on the 3G if I can and just whenever I get a convenient moment, to be honest. You can work it around yourself instead of having to physically go to a computer to do it...I prefer to use an app than going [and] logging into the website, personally.

Again, this view is reflected in their usage patterns—both spent a below average amount of time using the website; P10b did not use the website at all, whereas P9b only used the website for a short period of time during the first 2 weeks of the study. However, a strong preference for exclusive app-based access did not appear to be associated with longer or more frequent

use of the app. For example, P10b accessed the app on less than half the number of days it was available and spent an average of 3 minutes looking at the intervention content on each of those days, which is comparable to the time spent by other participants. P9b accessed the app less frequently and spent less time using the intervention content in comparison to other participants. A strong preference for exclusive app access was also not associated with stronger intervention effects. In fact, Access to POWer Tracker appeared to have a stronger effect on goal engagement for participants who valued using both the app and the website (P1a, P2a, P11b) than for participants who only valued the app (P9b, P10b).

When and Why Did Participants Use POWer Tracker?

Participants described using the app in short bursts during free moments of time throughout the day. On the days when the app was accessed, participants spent an average of 3 (SD 2 minutes) minutes per day using the intervention tools and 6 (SD 2 minutes) minutes per day completing the daily questionnaires. These short bursts of app use tended to be more prevalent during the morning (around 0900 and 1000 hours), lunchtime (around 1300 hours), and throughout the evening (starting from 1700 hours) (see [Figure 4](#)). For example, P11b said:

So really whenever you have free time. For me in between lectures you can go there and have a read...so that's quite an important part for me because sometimes you have 10 minute breaks so you can just go "Oh ok, I think I will take a look at my app and see what I can do."

There was variation between participants in when and how often the app was used within a 24-hour period. Most of the participants tended to access the app-based intervention content sporadically at different times throughout the day. For example, food lists were reportedly used at times when decisions were made about food choices, such as at the supermarket and while cooking or preparing meals. P10b described missing this point of reference when the app was not available:

When I was doing my food shopping I would usually use the card, the bit on it where it tells you which foods are good and which aren't, I have sort of based my food shopping around those lists, but without them this week I've had to just try and remember what was on there. So it was like a point of reference for me.

In contrast, a few participants only accessed the app-based intervention content in the morning or the evening. For example, when asked about when she used the POWer Tracker app, P4a answered:

Usually evening, so it's looking back on the day I mainly use it in the evening and then [I] can fill out the whole day...rather than bit by bit.

Of the 14 participants, 7 (54%) chose to set up automated notifications within the app (see [Table 4](#)). Only 2 participants discussed their reasons for not requesting notifications; one experienced a technical error preventing receipt of requested notifications (P3a) whereas the other reported that notifications were not needed because they could easily remember to use the

app without them (P10b). On visual inspection, participants who requested notifications tended to be those for whom access to the app had a stronger effect on goal engagement (eg, P11b, P8b, P1a, P12b), although there were exceptions (eg, P13b). However, no significant correlation was observed between the number of notifications received and effect of app access on goal engagement (awareness of eating goals: $r=-.07$, $P=.82$; achievement of eating goals: $r=-.17$, $P=.60$; awareness of physical activity goals: $r=.11$, $P=.75$).

Notifications were responded to 48 of 93 times (52%) they were received. On average, there was a 47-minute delay between notification receipt and subsequent response. View goal notifications were requested more often and prompted a greater and faster response (32/50, 64%; delay: mean 37, SD 44 minutes) than goal update notifications (16/43, 37%; delay: mean 67, SD 71 minutes). Responsiveness to notifications varied greatly both between and within individuals, ranging from the fastest response of 7 seconds and the slowest response of nearly 4 hours. The times at which notifications were requested also varied, including morning (0800 hours), lunchtime (1200 hours), early evening (1800-1900 hours), and late evening (2100-2300 hours). Participants commented that automated notifications were valuable for keeping goals in mind and prompting use of POWer. For example, P11b said:

The thing that partly remind[ed] me to go to the app [was] the reminders because sometimes you hear your phone sound and [it] remind[s] you of something...so you tend to look at your phone so maybe if the app [could] remind me to look at my phone [more] then I think it would be a bit better...once you take a look at the [message] I'm sure people [would] spend at least a couple of minutes taking a look at the app.

P11b also suggested that notifications may have more motivational impact if they contained personalized messages related to one's goals and plans before directing you to a particular component within the app.

Most participants explained that their app use was primarily determined by the availability of free moments within their daily routine, despite the usage prompts offered by time-relevant tools and automated notifications. Availability of free moments tended to be constrained by lifestyle factors such as work commitments and/or social commitments that were prioritized over using POWer. For example, P3a explains that automated notifications would only evoke a response if they were delivered at the right time:

If I was to receive something like [information on exercise or healthy eating] at 11 am when I was at work I would quite likely ignore it and never go back to it again. But say it came through in the evening when I was watching TV, things might be different and I might pay more attention to it.

The timing of unavailable moments tended to be more consistent than the timing of free moments within participants' daily routines. For example, P10b explained that she was nearly always unable to use the app during the day because she was busy with lectures, exams, and revision. However, her routine

varied as to whether she had free moments during the mornings, evenings, and weekend. Access to POWeR Tracker had the strongest effect on P1a, who also discussed how app use changed after a shift in daily routine (see Figure 5). During the first 2 weeks of the study, P1a was living at home with family while on summer vacation. He described using the app fairly infrequently and typically later in the evenings because he did not have a strict daily routine and also had limited control over meal choices:

It was just over summer so I wasn't really waking up that early. I guess if I had a 9-to-5 job then I would have used it in the mornings as well, but I wasn't waking up till about 11 or 12 because I'm off every day, by the time I had walked the dogs and done a few other things I guess I was going on the app then.

After returning to university and living in student accommodation, P1a reported using the app earlier in the day to take advantage of free Wi-Fi on campus as well as in the early hours of the morning while out with friends. P1a also expressed a greater need for the app when living independently:

I might start using it a bit more, because when I am at home obviously you have all your comforts...like bad foods that you shouldn't eat loads of biscuits around and stuff and so I was probably a little bit worse back home than I would be at university so I think maybe when I go shopping and stuff [now]...I'll have a quick look at the red, green, and yellow foods and simply won't buy the things so I don't have the temptation.

Figure 4. Average duration of participants' app use (minutes) by time of day for the first and second week of app access.

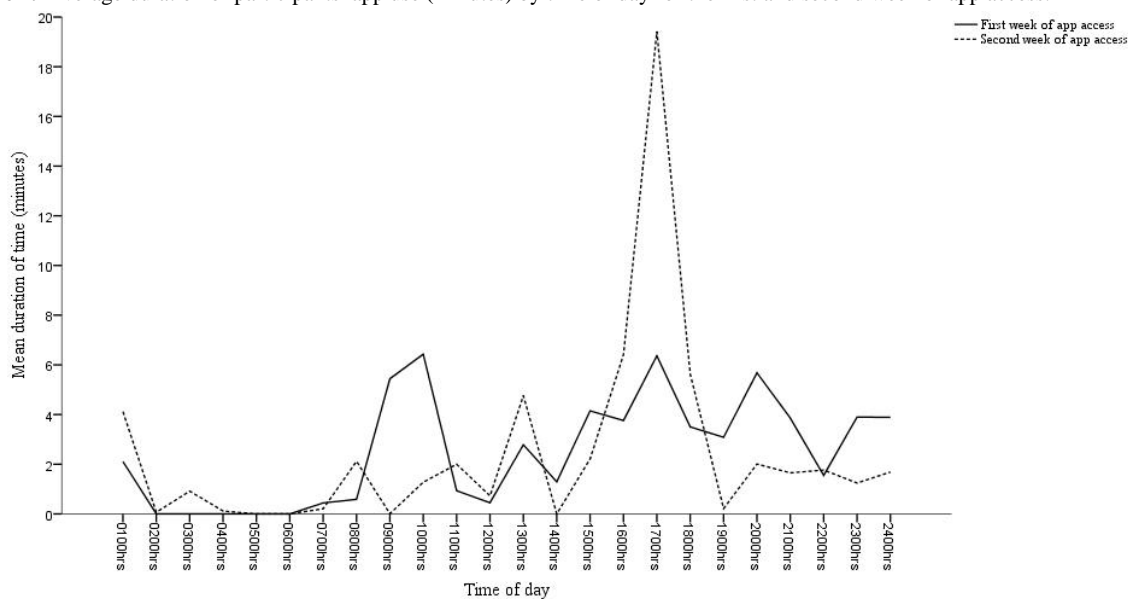
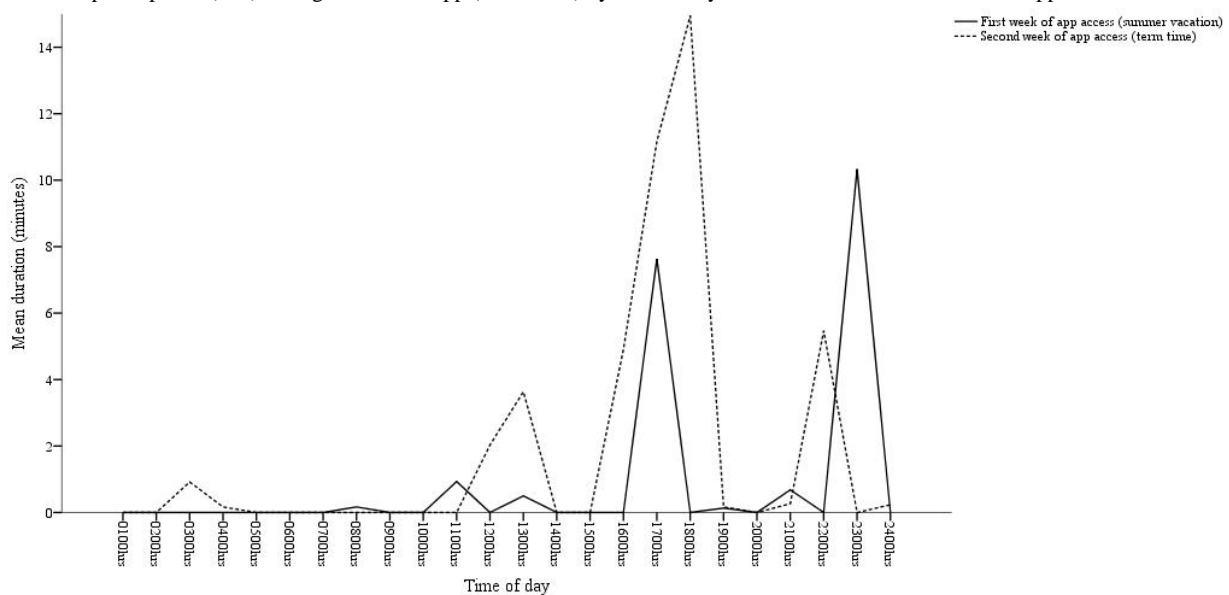


Figure 5. One participant's (P1a) average use of the app (in minutes) by time of day for the first and second weeks of app access.



Discussion

Principal Findings

Previous research has highlighted the benefits of stand-alone app-based interventions on individuals' awareness of their eating behaviors and physical activity levels [15,16,42,45,46]. The current study adds to this literature by demonstrating that provision of a hybrid Web- and app-based weight management intervention can encourage greater goal awareness than provision of a Web-based intervention alone. Overall, participants' awareness of both their eating and physical activity goals was greater during the weeks when the POWeR Tracker app was available. Qualitative data indicated that this effect appeared to be stronger for participants who valued the opportunity to access both Web- and app-based content as compared to participants who preferred to use the app exclusively. Usage data indicated that this effect also appeared to be stronger for participants who spent longer using the app and longer using the website during weeks when the app was not available. Access to the POWeR Tracker app was associated with improvement in participants' motivation, self-efficacy, and achievement of eating goals, but not physical activity goals. It is unsurprising that the POWeR Tracker app had greater impact on participants' engagement with their eating goals than physical activity goals given that the Web-based POWeR intervention places greater emphasis on changing day-to-day dietary routines. There is also growing evidence that individuals can struggle to pursue multiple goals and plans concurrently [64].

In-line with previous research [24,50,65], the POWeR Tracker app was viewed by most participants as a convenient means of accessing desired information and tools on-the-go as needed. However, the perceived convenience of POWeR Tracker did not necessarily deter participants from continuing to use the Web-based POWeR intervention, nor was it necessarily associated with regular use of the app and stronger effects of app access on self-reported goal engagement. Participants spent similar amounts of time using the POWeR website during the weeks when the app was and was not available indicating that the POWeR Tracker was used as a supplement, rather than replacement, for the Web-based intervention. Previous examinations of hybrid interventions designed to promote self-monitoring of behavior have reported that app-based delivery was associated with better adherence to self-monitoring compared to Web-based formats [17,44]. Unlike the interventions provided by Carter et al [17] and Kirwan et al [44], self-monitoring was not the only technique promoted by the POWeR intervention. Comparison of findings from these different types of intervention suggest that apps may be associated with greater usage when they are used to facilitate one specific, repetitive daily behavior, such as diary completion or step logging.

Participants in this study varied in their use of and preference for the informational versus self-monitoring components of the POWeR Tracker app. Although most participants did use POWeR Tracker to access informational content (eg, food lists), fewer used the app to track progress toward their goals or

complete food/physical activity diaries. Reasons cited for not using the tracking tools were the hassle of manual data entry and lack of personalized feedback. This contrasts with previous research revealing positive usage and views of app-based self-monitoring interventions [16,17,44], but does confirm findings from qualitative exploratory work suggesting that potential users of health apps are not always receptive to app-based self-monitoring [51,53]. As discussed earlier, self-monitoring is typically the sole behavioral target of studies that report positive usage and views of app-based delivery. It could be that these sorts of studies attract individuals who are already comfortable with self-monitoring their behavior and find it easier to integrate within their daily routines. It is also possible that the requirement for participants in the current study to complete daily measures via their mobile phones negated their perceived need for additional self-monitoring. Completion of the study measures accounted for the bulk of participants' usage of the POWeR Tracker app, with many participants commenting that answering daily measures was a useful motivational tool.

Participants' use of POWeR Tracker was not random and appeared to be triggered by 3 types of events: (1) relevance of app-based tools at particular times of day, (2) availability to respond or interact with the app, and (3) receipt of automated notifications. The impact of users' availability to interact with the app and take advantage of free moments echoes reports that ability to integrate app use into one's daily routine is an important potential facilitator and/or barrier to intervention usage [16,24,32,46]. There were no indications from participants in this study that automated reminder notifications from the app were perceived as annoying, a risk that has been highlighted previously [42]. This may be because the timing and receipt of all POWeR Tracker notifications were fully controlled by the participant, following guidelines suggested by Dennison et al [51].

Implications and Future Research

The findings from this study offer several implications for the future delivery of app-based health interventions. First, advances in mobile technology do not necessarily render Web-based interventions redundant. Combining app- and Web-based delivery in this study improved participants' self-reported engagement with their weight management goals suggesting that multicomponent, hybrid interventions may have the potential to enhance digital health promotion. Further research is needed to compare multicomponent, hybrid interventions to app-only interventions. This study also suggests that the benefit offered by app-based delivery in terms of convenience, does not necessarily correlate with usage or effect on outcomes, nor does it necessarily dissuade users from also engaging with Web-based interventions. That said, POWeR Tracker did not provide a complete duplication of the content and functionality offered by the Web-based intervention. Further research is necessary to examine whether the same usage patterns are observed when app-based components are equivalent to Web-based components.

Second, apps may be particularly well suited to the delivery of specific intervention components that are relevant on-the-go

and that can be accessed quickly during free moments. In the current study, app-based access to select informational content (eg, food lists) was viewed as particularly useful, whereas views of app-based self-monitoring were mixed. To date, considerable attention has been paid to whether existing, often commercial, apps provide theory and evidence-based BCTs [66-71]. An additional empirical question is whether particular BCTs are more or less suited to delivery using different digital formats (eg, app vs Web). For example, can all types of behavior change tools and techniques (eg, goal setting, cognitive behavior therapy approaches) feasibly be delivered via small mobile phone screens that may typically be accessed for only minutes at a time sporadically throughout the day? Are individuals willing to use their mobile phones for extended periods of time to access detailed intervention content? Does the future of app-based intervention lie as a supplemental component that delivers only the content that is useful and convenient to access on-the-go or as stand-alone interventions, or does this depend on factors such as the target behavior or intervention complexity? These are all questions that need to be addressed by future empirical research.

Third, this study points to the value of using a mixed-methods approach to understand how individuals use and view app-based delivery of health behavior change interventions. Analysis of only the qualitative data may have overemphasized the importance of participants' preferences for app-based delivery of intervention content. On the other hand, analysis of only the quantitative data would not have provided explanations for why the provision of a supplementary app had a greater impact for some participants compared to others. Furthermore, adopting an individualized approach enabled us to uncover differences in how participants reacted to and used the different behavior change tools provided by the POWeR Tracker app and whether these differences were related to the effect of app access on intervention outcomes. These insights will help us to improve the design of emerging app-based interventions by ensuring that they provide a range of different tools and techniques that have the potential to attract and retain a wider range of and/or higher numbers of engaged users. These insights also provide the warning that participants' preferences for intervention delivery (ie, apps are more convenient) may not necessarily match what actually leads to better outcomes (ie, combined use of website and app).

Building a detailed understanding of when and for what reasons individuals choose to use health apps is necessary to inform the development of intelligent systems that harness the phone's contextual sensor data to deliver real-time content or prompts at relevant and convenient moments. Adequate understanding of how and why users interact with app-based interventions on a day-to-day basis can also help to customize the design of interfaces that (1) enable users to easily access regularly used components of the app and (2) efficiently interact with components in a way that fits their actual often intermittent usage, rather than the regular usage researchers and/or programmers may intend [24,65]. Such intelligent and customized systems have the potential to both encourage and sustain usage of health-related apps.

Limitations

A number of limitations to this study should be noted. First, this study did not use an app- or Web-only control group or a true no-intervention control group. Therefore, no definitive conclusions can be drawn regarding the comparative impact of a stand-alone app versus Web-based delivery. We also cannot rule out the role of measurement effects. As previously discussed, the daily study measures were perceived by participants as an intervention tool that offered useful motivational benefits. Findings from this study are based on a small, predominantly young, and highly educated sample limiting generalizability and statistical power. Additional studies with larger and more diverse samples are required to confirm and replicate the findings observed in this study. Ceiling effects may have limited the potential impact of access to the POWeR Tracker app on certain individual's goal engagement, particularly goal awareness. Indeed, the participant who showed the strongest intervention effect (P1a) also had the lowest baseline values on day 1. All participants were motivated to learn about adopting a healthier lifestyle and some fell within the upper end of the healthy BMI range. It is unclear whether clinically overweight/obese individuals would interact with POWeR Tracker in the same way or whether providing access to the POWeR Tracker app would have the same level of impact on goal engagement.

It is also possible that participants' usage of both the Web-based POWeR intervention and POWeR Tracker app was confounded by aspects of the study design. For example, usage of the app during the evening may have been encouraged by the obligation to complete the daily study measures after 5 pm. It is also possible that the financial incentives offered for participation and the perceived obligation to prepare for the weekly telephone interviews may have led some participants in this study to use the Web-based intervention and app more regularly or in different ways than they otherwise might have. However, the fact that some participants did not use the website at all and/or used the app very infrequently indicates that this was not a concern for all participants. Additionally, there were a number of technical issues with the app software that limited how well we could explore participants' use and perceptions of the automated notifications provided by POWeR Tracker. The findings and implications regarding how participants interacted with a supplementary app-based tool provided alongside a Web-based intervention may also be specific to the domain of weight management.

Finally, this study measured the impact of providing access to POWeR Tracker on engagement with specific eating- and physical activity-related goals, but did not examine impact on dietary consumption or total physical activity level. Thus, we cannot know whether participants' increased awareness of their eating and physical activity goals translated into healthier dietary choices or increased activity levels. It is also not clear whether improving participants' experience of app-based interventions can actually lead to increased usage. For example, ratings of perceived usefulness and ease of use were not strongly associated with increased use of an app-based step logging tool [44].

Additional mixed-methods studies are needed to further examine both usage of and impact on behavioral change of providing supplemental app-based interventions in the longer term (ie, beyond 4 weeks). For example, does usage continue over the longer term and, if so, do and how do patterns of usage change? What factors influence long-term usage of health apps? Is there any association between a positive user experience and subsequent usage of health apps? Can exposure to health apps facilitate long-term maintenance of health behavior change?

Conclusions

Findings from this study suggest that supplementing a Web-based weight management intervention with an app-based

tool has the potential to improve individuals' motivation for and awareness of their healthy eating and physical activity goals. Perceived convenience of mobile-based access to intervention content enabled quick access to key pieces of intervention content on-the-go at relevant and convenient moments, but did not appear to deter use of the Web-based intervention. Using mixed-methods approaches can provide complementary qualitative and quantitative insights into how users view and use app-based health behavior interventions on a day-to-day basis and what impact app-based delivery may have on health-related goals.

Acknowledgments

We would like to thank all participants for sharing their views and experiences. We would also like to thank Hollie Butcher for her assistance with transcribing and cleaning the data. The study was funded by the UK Economic and Physical Sciences Research Council (EP/I032673/1: UBhave: ubiquitous and social computing for positive behavior change).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of content provided by the POWeR Tracker app.

[[PDF File \(Adobe PDF File\), 7KB - jmir_v16i10e201_app1.pdf](#)]

Multimedia Appendix 2

Interview schedule.

[[PDF File \(Adobe PDF File\), 5KB - jmir_v16i10e201_app2.pdf](#)]

Multimedia Appendix 3

Scale items for measures of goal awareness, achievement, self-efficacy and motivation.

[[PDF File \(Adobe PDF File\), 3KB - jmir_v16i10e201_app3.pdf](#)]

References

1. 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](#)]
2. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013;18(9):1039-1069. [doi: [10.1080/10810730.2013.768727](#)] [Medline: [23750972](#)]
3. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988-2007. *Prev Med* 2008 Jul;47(1):3-16 [FREE Full text] [doi: [10.1016/j.ypmed.2008.02.014](#)] [Medline: [18403003](#)]
4. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](#)] [Medline: [20164043](#)]
5. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME. "It's not that I don't have problems, I'm just not putting them on Facebook": challenges and opportunities in using online social networks for health. In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. 2011 Mar 19 Presented at: ACM 2011 Conference on Computer Supported Cooperative Work; Mar 19-23, 2011; Hangzhou, China p. 341-350. [doi: [10.1145/1958824.1958876](#)]
6. Munson S. Beyond the share Button: making social network sites work for health and wellness. *IEEE Potentials* 2011 Sep;30(5):42-47. [doi: [10.1109/MPOT.2011.941731](#)]

7. Buhi ER, Trudnak TE, Martinasek MP, Oberne AB, Fuhrmann HJ, McDermott RJ. Mobile phone-based behavioural interventions for health: A systematic review. *Health Education Journal* 2012 Jul 10;72(5):564-583. [doi: [10.1177/0017896912452071](https://doi.org/10.1177/0017896912452071)]
8. 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](https://doi.org/10.1016/j.amepre.2008.09.040)] [Medline: [19135907](https://pubmed.ncbi.nlm.nih.gov/19135907/)]
9. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. *J Biomed Inform* 2012 Feb;45(1):184-198 [FREE Full text] [doi: [10.1016/j.jbi.2011.08.017](https://doi.org/10.1016/j.jbi.2011.08.017)] [Medline: [21925288](https://pubmed.ncbi.nlm.nih.gov/21925288/)]
10. Miller G. The smartphone psychology manifesto. *Perspectives on Psychological Science* 2012 May 16;7(3):221-237. [doi: [10.1177/1745691612441215](https://doi.org/10.1177/1745691612441215)]
11. Morris ME, Aguilera A. Mobile, social, and wearable computing and the evolution of psychological practice. *Professional Psychology: Research and Practice* 2012;43(6):622-626. [doi: [10.1037/a0029041](https://doi.org/10.1037/a0029041)]
12. Hebden L, Cook A, van der Ploeg HP, Allman-Farinelli M. Development of smartphone applications for nutrition and physical activity behavior change. *JMIR Res Protoc* 2012;1(2):e9 [FREE Full text] [doi: [10.2196/resprot.2205](https://doi.org/10.2196/resprot.2205)] [Medline: [23611892](https://pubmed.ncbi.nlm.nih.gov/23611892/)]
13. Hong Y, Dahlke DV, Ory M, Hochhalter A, Reynolds J, Purcell NP, et al. Designing iCanFit: A Mobile-Enabled Web Application to Promote Physical Activity for Older Cancer Survivors. *JMIR Res Protoc* 2013;2(1):e12 [FREE Full text] [doi: [10.2196/resprot.2440](https://doi.org/10.2196/resprot.2440)] [Medline: [23612053](https://pubmed.ncbi.nlm.nih.gov/23612053/)]
14. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Design, development, and formative evaluation of a smartphone application for recording and monitoring physical activity levels: the 10,000 Steps "iStepLog". *Health Educ Behav* 2013 Apr;40(2):140-151. [doi: [10.1177/1090198112449460](https://doi.org/10.1177/1090198112449460)] [Medline: [22984196](https://pubmed.ncbi.nlm.nih.gov/22984196/)]
15. Klasnja P, Consolvo S, McDonald DW, Landay JA, Pratt W. Using mobile & personal sensing technologies to support health behavior change in everyday life: lessons learned. *AMIA Annu Symp Proc* 2009;2009:338-342 [FREE Full text] [Medline: [20351876](https://pubmed.ncbi.nlm.nih.gov/20351876/)]
16. Robinson E, Higgs S, Daley AJ, Jolly K, Lycett D, Lewis A, et al. Development and feasibility testing of a smart phone based attentive eating intervention. *BMC Public Health* 2013;13:639 [FREE Full text] [doi: [10.1186/1471-2458-13-639](https://doi.org/10.1186/1471-2458-13-639)] [Medline: [23837771](https://pubmed.ncbi.nlm.nih.gov/23837771/)]
17. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res* 2013;15(4):e32 [FREE Full text] [doi: [10.2196/jmir.2283](https://doi.org/10.2196/jmir.2283)] [Medline: [23587561](https://pubmed.ncbi.nlm.nih.gov/23587561/)]
18. Thomas JG, Wing RR. Health-e-call, a smartphone-assisted behavioral obesity treatment: pilot study. *JMIR Mhealth Uhealth* 2013 Apr;1(1):e3 [FREE Full text] [doi: [10.2196/mhealth.2164](https://doi.org/10.2196/mhealth.2164)] [Medline: [25100672](https://pubmed.ncbi.nlm.nih.gov/25100672/)]
19. Buller DB, Berwick M, Shane J, Kane I, Lantz K, Buller MK. User-centered development of a smart phone mobile application delivering personalized real-time advice on sun protection. *Transl Behav Med* 2013 Sep;3(3):326-334 [FREE Full text] [doi: [10.1007/s13142-013-0208-1](https://doi.org/10.1007/s13142-013-0208-1)] [Medline: [24058385](https://pubmed.ncbi.nlm.nih.gov/24058385/)]
20. McTavish FM, Chih MY, Shah D, Gustafson DH. How Patients Recovering From Alcoholism Use a Smartphone Intervention. *J Dual Diagn* 2012;8(4):294-304 [FREE Full text] [doi: [10.1080/15504263.2012.723312](https://doi.org/10.1080/15504263.2012.723312)] [Medline: [23316127](https://pubmed.ncbi.nlm.nih.gov/23316127/)]
21. Valdivieso-López E, Flores-Mateo G, Molina-Gómez JD, Rey-Reñones C, Barrera Uriarte ML, Duch J, et al. Efficacy of a mobile application for smoking cessation in young people: study protocol for a clustered, randomized trial. *BMC Public Health* 2013;13:704 [FREE Full text] [doi: [10.1186/1471-2458-13-704](https://doi.org/10.1186/1471-2458-13-704)] [Medline: [23915067](https://pubmed.ncbi.nlm.nih.gov/23915067/)]
22. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. *J Med Internet Res* 2012;14(3):e70 [FREE Full text] [doi: [10.2196/jmir.2058](https://doi.org/10.2196/jmir.2058)] [Medline: [22564332](https://pubmed.ncbi.nlm.nih.gov/22564332/)]
23. Katz R, Mesfin T, Barr K. Lessons from a community-based mHealth diabetes self-management program: "it's not just about the cell phone". *J Health Commun* 2012;17 Suppl 1:67-72. [doi: [10.1080/10810730.2012.650613](https://doi.org/10.1080/10810730.2012.650613)] [Medline: [22548601](https://pubmed.ncbi.nlm.nih.gov/22548601/)]
24. Tataru N, Arsand E, Skrøvseth SO, Hartvigsen G. Long-term engagement with a mobile self-management system for people with type 2 diabetes. *JMIR Mhealth Uhealth* 2013;1(1):e1 [FREE Full text] [doi: [10.2196/mhealth.2432](https://doi.org/10.2196/mhealth.2432)] [Medline: [25100649](https://pubmed.ncbi.nlm.nih.gov/25100649/)]
25. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. *J Med Internet Res* 2013;15(1):e5 [FREE Full text] [doi: [10.2196/jmir.2249](https://doi.org/10.2196/jmir.2249)] [Medline: [23291270](https://pubmed.ncbi.nlm.nih.gov/23291270/)]
26. Marcano Belisario JS, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. *Cochrane Database Syst Rev* 2013;11:CD010013. [doi: [10.1002/14651858.CD010013.pub2](https://doi.org/10.1002/14651858.CD010013.pub2)] [Medline: [24282112](https://pubmed.ncbi.nlm.nih.gov/24282112/)]
27. Donker T, Petrie K, Proudfoot J, Clarke J, Birch MR, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res* 2013;15(11):e247 [FREE Full text] [doi: [10.2196/jmir.2791](https://doi.org/10.2196/jmir.2791)] [Medline: [24240579](https://pubmed.ncbi.nlm.nih.gov/24240579/)]
28. Stegemann SK, Ebenfeld L, Lehr D, Berking M, Funk B. Development of a mobile application for people with panic disorder as augmentation for an Internet-based intervention. In: *Proceedings of the Federated Conference on Computer Science and Information Systems*. 2013 Sep 08 Presented at: Federated Conference on Computer Science and Information Systems; Sep 8-11, 2013; Kraków, Poland p. 1319-1325.

29. Lindner P, Ivanova E, Ly KH, Andersson G, Carlbring P. Guided and unguided CBT for social anxiety disorder and/or panic disorder via the Internet and a smartphone application: study protocol for a randomised controlled trial. *Trials* 2013;14:437 [FREE Full text] [doi: [10.1186/1745-6215-14-437](https://doi.org/10.1186/1745-6215-14-437)] [Medline: [24351088](https://pubmed.ncbi.nlm.nih.gov/24351088/)]
30. Ahtinen A, Huuskonen P. Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending boundaries. In: Let's all get up and walk to the North Pole: design and evaluation of a mobile wellness application. 2010 Oct 16 Presented at: 6th Nordic Conference on Human-Computer Interaction: Extending boundaries; Oct 16 - 20, 2010; Reykjavik, Iceland p. 3-12. [doi: [10.1145/1868914.1868920](https://doi.org/10.1145/1868914.1868920)]
31. Lane N, Mohammad M, Lin M, Yang X, Lu H, Ali S, et al. BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In: Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare. 2011 May Presented at: 5th International ICST Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2011; Dublin, Republic of Ireland p. 23-36. [doi: [10.4108/icst.pervasivehealth.2011.246161](https://doi.org/10.4108/icst.pervasivehealth.2011.246161)]
32. Ly KH, Dahl J, Carlbring P, Andersson G. Development and initial evaluation of a smartphone application based on acceptance and commitment therapy. *Springerplus* 2012;1:11. [doi: [10.1186/2193-1801-1-11](https://doi.org/10.1186/2193-1801-1-11)] [Medline: [23961342](https://pubmed.ncbi.nlm.nih.gov/23961342/)]
33. Kuhn E, Greene C, Hoffman J, Nguyen T, Wald L, Schmidt J, et al. Preliminary evaluation of PTSD Coach, a smartphone app for post-traumatic stress symptoms. *Mil Med* 2014 Jan;179(1):12-18. [doi: [10.7205/MILMED-D-13-00271](https://doi.org/10.7205/MILMED-D-13-00271)] [Medline: [24402979](https://pubmed.ncbi.nlm.nih.gov/24402979/)]
34. Schmer-Galunder S, Keller P. Remote anger management support using mobile technology - an android phone application for persons with post traumatic stress disorder (PTSD). In: SOTICS 2013. 2013 Nov 17 Presented at: The Third International Conference on Social Eco-Informatics; Nov 17-22, 2013; Lisbon, Portugal p. 61-66.
35. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. *J Med Internet Res* 2011;13(3):e55 [FREE Full text] [doi: [10.2196/jmir.1838](https://doi.org/10.2196/jmir.1838)] [Medline: [21840837](https://pubmed.ncbi.nlm.nih.gov/21840837/)]
36. Bardram JE, Frost M, Marcu G. The MONARCA self-assessment system – a persuasive personal monitoring system for bipolar patients. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium. 2012 Jan 28 Presented at: 2nd ACM SIGHIT International Health Informatics Symposium; Jan 28-30, 2012; Miami, FL p. 21. [doi: [10.1145/2110363.2110370](https://doi.org/10.1145/2110363.2110370)]
37. Wenze SJ, Arney MF, Miller IW. Feasibility and Acceptability of a Mobile Intervention to Improve Treatment Adherence in Bipolar Disorder: A Pilot Study. *Behav Modif* 2014 Jan 8;38(4):497-515. [doi: [10.1177/0145445513518421](https://doi.org/10.1177/0145445513518421)] [Medline: [24402464](https://pubmed.ncbi.nlm.nih.gov/24402464/)]
38. Morgan PJ, Lubans DR, Collins CE, Warren JM, Callister R. The SHED-IT randomized controlled trial: evaluation of an Internet-based weight-loss program for men. *Obesity (Silver Spring)* 2009 Nov;17(11):2025-2032. [doi: [10.1038/oby.2009.85](https://doi.org/10.1038/oby.2009.85)] [Medline: [19343018](https://pubmed.ncbi.nlm.nih.gov/19343018/)]
39. Løventoft PK, Nørregaard LB, Frøkjær E, København S. Designing Daybuilder: an experimental app to support people with depression. In: Proceedings of the 12th Participatory Design Conference. 2012 Aug 12 Presented at: 12th Participatory Design Conference; Aug 12-16, 2012; Roskilde, Denmark p. 12-16. [doi: [10.1145/2348144.2348146](https://doi.org/10.1145/2348144.2348146)]
40. Ben-Zeev D, Kaiser SM, Brenner CJ, Begale M, Duffecy J, Mohr DC. Development and usability testing of FOCUS: a smartphone system for self-management of schizophrenia. *Psychiatr Rehabil J* 2013 Dec;36(4):289-296. [doi: [10.1037/prj0000019](https://doi.org/10.1037/prj0000019)] [Medline: [24015913](https://pubmed.ncbi.nlm.nih.gov/24015913/)]
41. Lister C, West JH, Richards R, Crookston B, Hall PC, Redelfs AH. Technology for health: A qualitative study on barriers to using the iPad for diet change. *Health* 2013;05(04):761-768. [doi: [10.4236/health.2013.54101](https://doi.org/10.4236/health.2013.54101)]
42. Fukuoka Y, Lindgren T, Jong S. Qualitative exploration of the acceptability of a mobile phone and pedometer-based physical activity program in a diverse sample of sedentary women. *Public Health Nurs* 2012;29(3):232-240. [doi: [10.1111/j.1525-1446.2011.00997.x](https://doi.org/10.1111/j.1525-1446.2011.00997.x)] [Medline: [22512424](https://pubmed.ncbi.nlm.nih.gov/22512424/)]
43. Rabin C, Bock B. Desired features of smartphone applications promoting physical activity. *Telemed J E Health* 2011 Dec;17(10):801-803. [doi: [10.1089/tmj.2011.0055](https://doi.org/10.1089/tmj.2011.0055)] [Medline: [22010977](https://pubmed.ncbi.nlm.nih.gov/22010977/)]
44. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. *J Med Internet Res* 2012;14(2):e55 [FREE Full text] [doi: [10.2196/jmir.1950](https://doi.org/10.2196/jmir.1950)] [Medline: [22522112](https://pubmed.ncbi.nlm.nih.gov/22522112/)]
45. Carroll EA, Czerwinski M, Roseway A, Kapoor A, Johns P, Rowan K, et al. Food and Mood: Just-in-Time Support for Emotional Eating. 2013 Sep 02 Presented at: Humaine Association Conference on Affective Computing and Intelligent Interaction; Sep 2-5, 2013; Geneva, Switzerland p. 252-257. [doi: [10.1109/ACII.2013.48](https://doi.org/10.1109/ACII.2013.48)]
46. Anderson I, Maitland J, Sherwood S, Barkhuus L, Chalmers M, Hall M, et al. Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Netw Appl* 2007 Aug 3;12(2-3):185-199. [doi: [10.1007/s11036-007-0011-7](https://doi.org/10.1007/s11036-007-0011-7)]
47. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. *J Med Internet Res* 2013;15(5):e95 [FREE Full text] [doi: [10.2196/jmir.2430](https://doi.org/10.2196/jmir.2430)] [Medline: [23697600](https://pubmed.ncbi.nlm.nih.gov/23697600/)]
48. Ozdalga E, Ozdalga A, Ahuja N. The smartphone in medicine: a review of current and potential use among physicians and students. *J Med Internet Res* 2012;14(5):e128 [FREE Full text] [doi: [10.2196/jmir.1994](https://doi.org/10.2196/jmir.1994)] [Medline: [23017375](https://pubmed.ncbi.nlm.nih.gov/23017375/)]
49. Pulman A, Taylor J, Galvin K, Masding M. Ideas and enhancements related to mobile applications to support type 1 diabetes. *JMIR Mhealth Uhealth* 2013 Jul;1(2):e12 [FREE Full text] [doi: [10.2196/mhealth.2567](https://doi.org/10.2196/mhealth.2567)] [Medline: [25100684](https://pubmed.ncbi.nlm.nih.gov/25100684/)]

50. Seto E, Leonard KJ, Masino C, Cafazzo JA, Barnsley J, Ross HJ. Attitudes of heart failure patients and health care providers towards mobile phone-based remote monitoring. *J Med Internet Res* 2010;12(4):e55 [FREE Full text] [doi: [10.2196/jmir.1627](https://doi.org/10.2196/jmir.1627)] [Medline: [21115435](https://pubmed.ncbi.nlm.nih.gov/21115435/)]
51. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
52. Fukuoka Y, Kamitani E, Bonnet K, Lindgren T. Real-time social support through a mobile virtual community to improve healthy behavior in overweight and sedentary adults: a focus group analysis. *J Med Internet Res* 2011;13(3):e49 [FREE Full text] [doi: [10.2196/jmir.1770](https://doi.org/10.2196/jmir.1770)] [Medline: [21752785](https://pubmed.ncbi.nlm.nih.gov/21752785/)]
53. Vandelanotte C, Caperchione CM, Ellison M, George ES, Maeder A, Kolt GS, et al. What kinds of website and mobile phone-delivered physical activity and nutrition interventions do middle-aged men want? *J Health Commun* 2013;18(9):1070-1083. [doi: [10.1080/10810730.2013.768731](https://doi.org/10.1080/10810730.2013.768731)] [Medline: [23647448](https://pubmed.ncbi.nlm.nih.gov/23647448/)]
54. Sultan S, Mohan P. Transforming usage data into a sustainable mobile health solution. *Electron Markets* 2012 Apr 29;23(1):63-72. [doi: [10.1007/s12525-012-0090-6](https://doi.org/10.1007/s12525-012-0090-6)]
55. Schwarzer R. Modeling Health Behavior Change: How to Predict and Modify the Adoption and Maintenance of Health Behaviors. *Applied Psychology* 2008 Jan;57(1):1-29. [doi: [10.1111/j.1464-0597.2007.00325.x](https://doi.org/10.1111/j.1464-0597.2007.00325.x)]
56. Sniehotta FF, Scholz U, Schwarzer R. Bridging the intention-behaviour gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health* 2005 Apr;20(2):143-160. [doi: [10.1080/08870440512331317670](https://doi.org/10.1080/08870440512331317670)]
57. Yardley L, Williams S, Bradbury K, Garip G, Renouf S, Ware L, et al. Integrating user perspectives into the development of a web-based weight management intervention. *Clinical Obesity* 2013 Jan 14;2(5-6):132-141. [doi: [10.1111/cob.12001](https://doi.org/10.1111/cob.12001)]
58. Yardley L, Ware LJ, Smith ER, Williams S, Bradbury KJ, Arden-Close EJ, et al. Randomised controlled feasibility trial of a web-based weight management intervention with nurse support for obese patients in primary care. *Int J Behav Nutr Phys Act* 2014;11:67 [FREE Full text] [doi: [10.1186/1479-5868-11-67](https://doi.org/10.1186/1479-5868-11-67)] [Medline: [24886516](https://pubmed.ncbi.nlm.nih.gov/24886516/)]
59. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial. *J Med Internet Res* 2014;16(3):e95 [FREE Full text] [doi: [10.2196/jmir.3199](https://doi.org/10.2196/jmir.3199)] [Medline: [24681761](https://pubmed.ncbi.nlm.nih.gov/24681761/)]
60. University of Southampton. 2014. LifeGuide Community Website URL: <https://www.lifeguideonline.org/> [accessed 2014-05-13] [WebCite Cache ID 6PXFImq4p]
61. Louro MJ, Pieters R, Zeelenberg M. Dynamics of multiple-goal pursuit. *J Pers Soc Psychol* 2007 Aug;93(2):174-193. [doi: [10.1037/0022-3514.93.2.174](https://doi.org/10.1037/0022-3514.93.2.174)] [Medline: [17645394](https://pubmed.ncbi.nlm.nih.gov/17645394/)]
62. R Development Core Team. A language and environment for statistical computing. 2008. URL: <http://www.r-project.org/> [accessed 2014-05-12] [WebCite Cache ID 6PVt6OVXX]
63. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
64. Verhoeven AAC, Adriaanse MA, de Ridder DTD, de Vet E, Fennis BM. Less is more: The effect of multiple implementation intentions targeting unhealthy snacking habits. *Eur J Soc Psychol* 2013 Jul 02;43(5):344-354. [doi: [10.1002/ejsp.1963](https://doi.org/10.1002/ejsp.1963)]
65. Mattila E, Orsama AL, Ahtinen A, Hopsu L, Leino T, Korhonen I. Personal health technologies in employee health promotion: usage activity, usefulness, and health-related outcomes in a 1-year randomized controlled trial. *JMIR Mhealth Uhealth* 2013 Jul;1(2):e16 [FREE Full text] [doi: [10.2196/mhealth.2557](https://doi.org/10.2196/mhealth.2557)] [Medline: [25098385](https://pubmed.ncbi.nlm.nih.gov/25098385/)]
66. Breton ER, Fuemmeler BF, Abrams 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](https://doi.org/10.1007/s13142-011-0076-5)] [Medline: [24073074](https://pubmed.ncbi.nlm.nih.gov/24073074/)]
67. 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](https://doi.org/10.2196/jmir.1977)] [Medline: [22584372](https://pubmed.ncbi.nlm.nih.gov/22584372/)]
68. 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](https://doi.org/10.1177/1090198112452126)] [Medline: [22991048](https://pubmed.ncbi.nlm.nih.gov/22991048/)]
69. 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](https://doi.org/10.1007/s13142-013-0205-4)] [Medline: [24073179](https://pubmed.ncbi.nlm.nih.gov/24073179/)]
70. 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](https://doi.org/10.1016/j.amepre.2013.07.005)] [Medline: [24139771](https://pubmed.ncbi.nlm.nih.gov/24139771/)]
71. Pagoto S, Schneider K, Jovic 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](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]

Abbreviations

BCT: behavior change technique

BMI: body mass index

POWeR: Positive Online Weight Reduction

Edited by G Eysenbach; submitted 04.06.14; peer-reviewed by S Winter, M Bardus, C Maher, A Middelweerd; comments to author 25.06.14; revised version received 01.08.14; accepted 20.08.14; published 22.10.14.

Please cite as:

Morrison LG, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, Michaelides DT, Johnston D, Johnston M, Michie S, Little P, Smith PWF, Weal MJ, Yardley L

Understanding Usage of a Hybrid Website and Smartphone App for Weight Management: A Mixed-Methods Study

J Med Internet Res 2014;16(10):e201

URL: <http://www.jmir.org/2014/10/e201/>

doi: [10.2196/jmir.3579](https://doi.org/10.2196/jmir.3579)

PMID: [25355131](https://pubmed.ncbi.nlm.nih.gov/25355131/)

©Leanne G Morrison, Charlie Hargood, Sharon Xiaowen Lin, Laura Dennison, Judith Joseph, Stephanie Hughes, Danius T Michaelides, Derek Johnston, Marie Johnston, Susan Michie, Paul Little, Peter WF Smith, Mark J Weal, Lucy Yardley. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 22.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Diabetes Patients' Experiences With the Implementation of Insulin Therapy and Their Perceptions of Computer-Assisted Self-Management Systems for Insulin Therapy

Airin CR Simon¹, MD; Wouter T Gude², BSc (Hons); Frits Holleman³, MD, PhD; Joost BL Hoekstra³, MD, PhD; Niels Peek^{2,4}, PhD

¹University of Amsterdam/Academic Medical Center, Internal Medicine/Medical Informatics, Amsterdam, Netherlands

²Academic Medical Center, Medical Informatics, University of Amsterdam, Amsterdam, Netherlands

³Academic Medical Center, Internal Medicine, University of Amsterdam, Amsterdam, Netherlands

⁴Health e-Research Centre, The University of Manchester, Manchester, United Kingdom

Corresponding Author:

Airin CR Simon, MD

University of Amsterdam/Academic Medical Center

Internal Medicine/Medical Informatics

Rm F4-215

Meibergdreef 9

Amsterdam, 1105 AZ

Netherlands

Phone: 31 205668136

Fax: 31 206914904

Email: a.c.simon@amc.uva.nl

Abstract

Background: Computer-assisted decision support is an emerging modality to assist patients with type 2 diabetes mellitus (T2DM) in insulin self-titration (ie, self-adjusting insulin dose according to daily blood glucose levels). Computer-assisted insulin self-titration systems mainly focus on helping patients overcome barriers related to the cognitive components of insulin titration. Yet other (eg, psychological or physical) barriers could still impede effective use of such systems.

Objective: Our primary aim was to identify experiences with and barriers to self-monitoring of blood glucose, insulin injection, and insulin titration among patients with T2DM. Our research team developed a computer-assisted insulin self-titration system, called PANDIT. The secondary aim of this study was to evaluate patients' perceptions of computer-assisted insulin self-titration. We included patients who used PANDIT in a 4-week pilot study as well as patients who had never used such a system.

Methods: In-depth, semi-structured interviews were conducted individually with patients on insulin therapy who were randomly recruited from a university hospital and surrounding general practices in the Netherlands. The interviews were transcribed verbatim and analyzed qualitatively. To classify the textual remarks, we created a codebook during the analysis, in a bottom-up and iterative fashion. To support examination of the final coded data, we used three theories from the field of health psychology and the integrated model of user satisfaction and technology acceptance by Wixom and Todd.

Results: When starting insulin therapy, some patients feared a lifelong commitment to insulin therapy and disease progression. Also, many barriers arose when implementing insulin therapy (eg, some patients were embarrassed to inject insulin in public). Furthermore, patients had difficulties increasing the insulin dose because they fear hypoglycemia, they associate higher insulin doses with disease progression, and some were ignorant of treatment targets. Patients who never used a computer-assisted insulin self-titration system felt they had enough knowledge to know when their insulin should be adjusted, but still believed that the system advice would be useful to confirm their reasoning. Furthermore, the time and effort saved with automated insulin advice was considered an advantage. Patients who had used PANDIT found the system useful if their glycemic regulation improved. Nevertheless, for some patients, the absence of personal contact with their caregiver was a drawback. While guidelines state that adjustment of basal insulin dose based on fasting plasma glucose values is sufficient, both patients who had and those who had not used PANDIT felt that such a system should take more patient data into consideration, such as lifestyle and diet factors.

Conclusions: Patients encounter multiple obstacles when implementing insulin therapy. Computer-assisted insulin self-titration can increase patient awareness of treatment targets and increase their confidence in self-adjusting the insulin dose. Nevertheless, some barriers may still exist when using computer-assisted titration systems and these systems could also introduce new barriers.

(*J Med Internet Res* 2014;16(10):e235) doi:[10.2196/jmir.3198](https://doi.org/10.2196/jmir.3198)

KEYWORDS

Type 2 diabetes mellitus; clinical decision support systems; patient acceptance of health care

Introduction

The rising incidence of type 2 diabetes mellitus (T2DM) is a major public health concern and financial burden [1]. Complications associated with this disease can be reduced by lowering blood glucose levels to near normal values with diet, oral glucose lowering treatment, and/or insulin therapy. For many patients, diet and oral glucose lowering treatment are insufficient to maintain good glycemic control in the long term. The progressive deterioration of the beta-cell function necessitates the addition of insulin to their existing treatment. Many clinical trials have shown that good glycemic control can be achieved in the majority of patients on insulin therapy [2]. International guidelines recommend a glycemic target of HbA1c (glycated hemoglobin) <7% (53 mmol/mol) to reduce diabetes-related complications [3,4]. In the United Kingdom and in the United States, only 59% and 47% of patients with T2DM achieve HbA1c values of 7.4% and 6-8% [5,6]. The glycemic target is also not achieved in most patients with T2DM in the Netherlands [7].

Insulin therapy requires frequent self-monitoring of blood glucose (SMBG), that is, observing and recording daily blood glucose levels over time to facilitate accordant adjustment of the insulin dose, called insulin titration [8]. While SMBG and injection of insulin are performed mostly at the patients' homes [9], the majority of patients with T2DM requiring insulin therapy have their doses titrated by their care providers [10]. Failures to reach treatment target in clinical practice are partly due to the fact that patients have difficulties implementing insulin therapy; a nationwide study in the United States showed that 29% of patients with T2DM treated with insulin never practiced SMBG [11]. A systematic review published in 2004 reported that adherence rates to insulin varied from 62% to 64% in patients with T2DM in developed countries [12]. Furthermore, evidence suggests that patients remain on low doses of insulin and that their insulin doses are adjusted insufficiently to achieve treatment targets [3].

Several studies have shown that self-management is a key component of effective care for chronic diseases, such as T2DM, leading to improved patient outcomes [13]. If insulin titration can also be undertaken successfully by patients with T2DM themselves, this might improve glycemic control. Among patients with type 1 diabetes mellitus (T1DM), insulin self-titration is already well-established suggesting that therapeutic self-management may be similarly beneficial when applied to patients with T2DM [14].

Computer-assisted decision support is an emerging modality to assist patients with chronic diseases in general and patients with

diabetes in the self-management of their disease [15]. Many existing systems in this area focus on insulin self-titration to support patients with T1DM in calculating the optimal pre-meal short-acting insulin dose [16]. Patients with T2DM who become insulin-dependent start with a basal insulin and thus require a different titration strategy [17]. To date, the majority of systems for patients with T2DM allow patients to enter blood glucose values in an electronic diary and are complemented by telemedicine functionalities to receive recommendations from medical professionals after the glucose readings are assessed [18,19]. More advanced systems automatically provide insulin dosing advice through decision support technologies [20]. While such computer-assisted insulin self-titration systems focus mainly on helping patients overcome barriers to the cognitive components of insulin titration, there may be other barriers, such as psychological or physical barriers, that still impede effective use of such systems. Furthermore, the use of computer-assisted systems by patients could induce new barriers during their self-management.

Many studies have focused on caregivers' and patients' resistance to the initiation of insulin therapy [21-23]. Studies have also identified barriers to the evaluation of blood glucose levels [24-27] and to insulin injection [12,28,29]. Yet, to our knowledge, no study has focused on patients' experiences with and barriers to insulin titration. Also, no study has examined patients' perceptions of computer-assisted insulin self-titration, investigating the feasibility of these decision support systems to be implemented in the near future. Our research group developed a computer-assisted insulin self-titration system for patients using once daily basal insulin, called the Patient Assisting Net-based Diabetes Insulin Titration (PANDIT) system [30]. In this qualitative study, our primary aim was therefore to identify experiences with and barriers to SMBG, insulin injection, and insulin titration among patients with T2DM using in-depth, semi-structured interviews. These experiences and barriers to insulin self-management were analyzed using three theories from the field of health psychology. The secondary aim of this study was to evaluate the perceptions of patients who used PANDIT in a pilot study as well as patients who never used such a system. This second analysis was based on the integrated model of user satisfaction and technology acceptance by Wixom and Todd [31].

Methods

Study Participants

Patients with T2DM, aged 18-80 years, who were on insulin therapy, were considered eligible to participate in this study, irrespective of their glycemic control. In this study, we aimed to include 10 participants who had never used a

computer-assisted insulin self-titration system (Group 1). Furthermore, 10 patients who were enrolled in a 4-week pilot involving the use of PANDIT were asked to participate in this study (Group 2). All participants were randomly recruited from a 1000-bed university hospital in Amsterdam, and general practices in and around Amsterdam, the Netherlands.

Description of PANDIT

PANDIT is a Web-based system that provides insulin dosing advice for patients using once daily basal insulin. Exogenous basal insulin complements the residual insulin that is supplied by the pancreas. Basal insulin doses are traditionally titrated according to fasting plasma glucose (FPG) values [32].

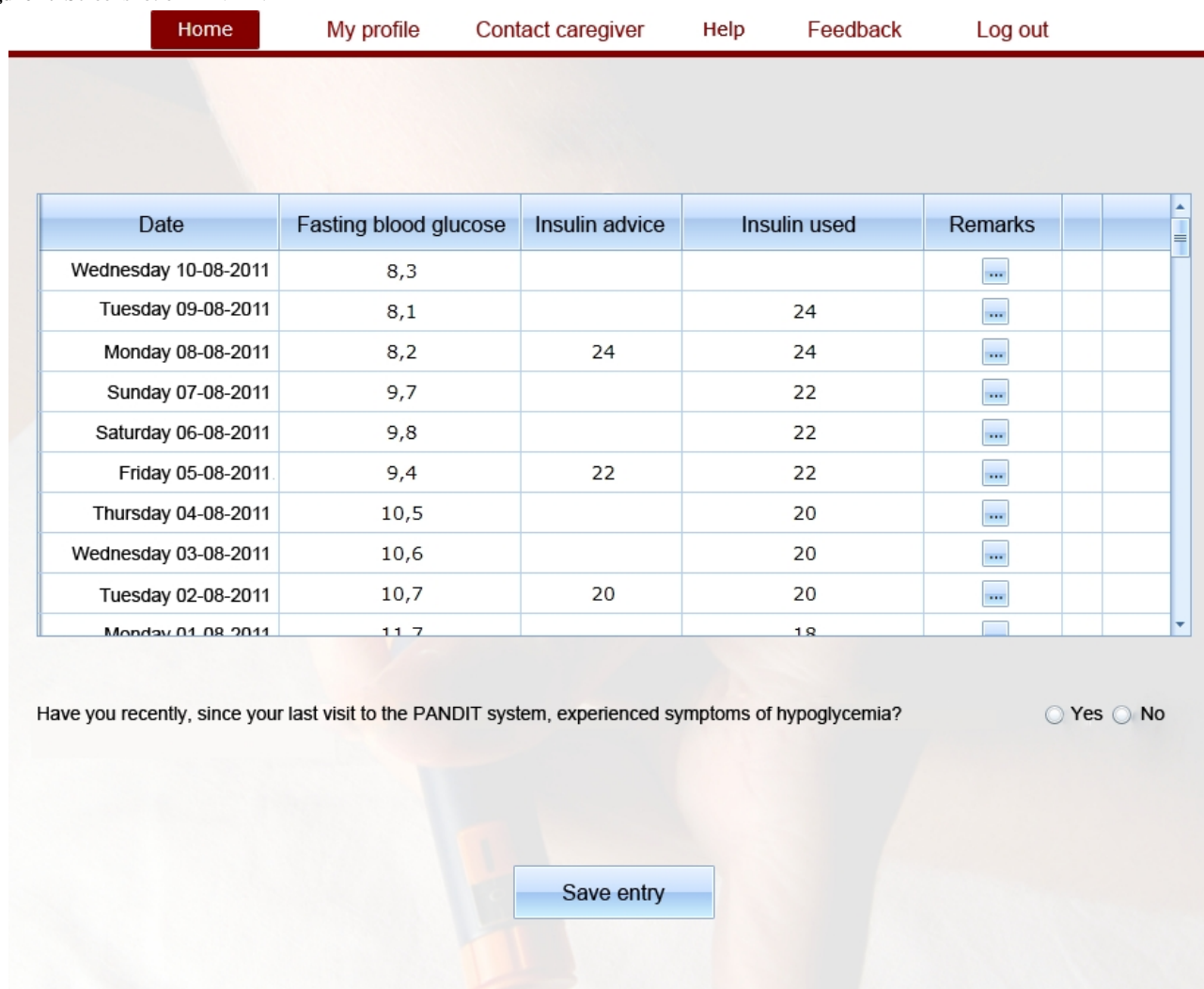
The patient user interface of PANDIT resembles a glucose diary in order to facilitate the collection of FPG values. The patient user interface of PANDIT is displayed in Figure 1. After a patient has logged in and opened the diary, a one-page screen is displayed containing five columns that show calendar date, FPG values, insulin dosing advice as provided by the system, current dose of insulin used, and remarks. Patients need to access PANDIT at least once every 3 days to enter recently measured FPG values and their current insulin dose. Furthermore, they

have to indicate whether they have experienced symptoms of hypoglycemic episodes.

A decision support algorithm uses the measured FPG values to determine whether the patient’s FPG is within the target range. If this is the case, the patient is advised to continue the current insulin dose, otherwise, the system will advise to adjust the insulin dose. If there is a reason for more intensively guided treatment, for example, if the patient frequently experiences hypoglycemic episodes, PANDIT will be automatically blocked from generating new insulin dosing advice and the care provider receives a message from PANDIT.

In addition to the decision support algorithm, PANDIT also incorporates asynchronous telemedicine functionalities. The telemedicine functionalities allow care providers to provide insulin dosing advice through the PANDIT interface to their patients. These telemedicine functionalities are automatically triggered when the system is blocked from generating advice as described above, but can also be evoked by care providers when they think this is necessary. As soon as the patient is sufficiently stabilized, the care provider can decide to “unblock” the algorithm, and let PANDIT generate new insulin dosing advice again.

Figure 1. Screenshot of PANDIT.



Data Collection

In-depth, semi-structured interviews were conducted individually with all participants of the study. One researcher (ACRS) conducted face-to-face interviews with the participants of Group 1 to investigate their experiences with and barriers to SMBG, insulin injection, and insulin titration, as well as their perceptions of computer-assisted insulin self-titration. The face-to-face interviews were conducted at the outpatient clinic.

Another researcher (WTG) conducted telephone interviews with the participants of Group 2 specifically focusing on their experiences with PANDIT. Face-to-face interviews had an average duration of 60 minutes. Telephone interviews lasted around 20 minutes. All interviews were audio-taped with oral consent of the participants. Table 1 lists the topics that were addressed during the interviews. Open-ended questions with neutral probes related to each topic were posed and follow-up questions were used to elicit more in-depth information.

Table 1. List of topics that were addressed during the interviews.

Topic	Item
Face-to-face interviews (Group 1)	
A. Experiences with and barriers to insulin titration	
To what extent patients adhere to...(a-c)	(a) performing SMBG ^a
How patients feel about...(a-c)	(b) injection of insulin
The ideal conditions to...(a-c)	(c) titration of insulin
Difficult conditions to...(a-c)	
B. Perceptions of computer-assisted insulin self-titration	
To what extent patients are skillful in using...(a,b)	(a) computer
How patients feel about using...(a-c)	(b) telemedicine
	(c) computer-assisted insulin self-titration system
Telephone interviews (Group 2)	
Experiences with and barriers to the use of PANDIT^b	
Positive or negative aspects/functionalities of PANDIT	
Missing aspects/functionalities of PANDIT	
Feelings about the insulin dosing advice given through PANDIT	

^aSMBG: self-monitoring of blood glucose

^bPANDIT: Patient Assisting Net-based Diabetes Insulin Titration system

Data Analysis

The interviews were transcribed verbatim and analyzed qualitatively. In the analysis, we used the MAXQDA software [33]. For the primary aim, two researchers (ACRS and WTG) individually extracted all remarks regarding patients' experiences with and barriers to SMBG, insulin injection, and insulin titration from the face-to-face interviews with Group 1. For the secondary aim, the two researchers extracted the perceptions of computer-assisted insulin self-titration addressed by participants in Group 1 and the experiences with PANDIT addressed by participants in Group 2.

Analysis of Experiences With and Barriers to Insulin Self-Management

The analysis of the interview data was based on the steps described in the handbook of Green et al [34]. To classify the textual remarks, we created a codebook during the analysis, in a bottom-up and iterative fashion, following the procedure for codebook creation described by DeCuir-Gunby et al [35]. The codes were developed from the data rather than from existing theories, and we examined the data multiple times and revised the codebook and the assigned codes if necessary, until there

were no more changes either to the codebook or the assigned codes. During the second and later cycles of examination, we used constructs from three selected theories from the field of health psychology (Stages of Change Model [36], Theory of Planned Behavior [37], and Self-Regulatory Theory [38]) to guide changes to the codebook. However, data was never "forced" into these theories. If no suitable construct was found, the original code that was derived from the data was retained.

Figure 2 depicts the process that was followed to match bottom-up coded remarks to constructs from the three theories from health psychology. The Stages of Change Model of Prochaska and DiClemente distinguishes five stages on the way to behavioral change, namely precontemplation, contemplation, preparation, action, and maintenance [24]. The Stages of Change Model hence was used to appreciate patients' readiness to perform insulin self-management, requiring acknowledgement of starting insulin therapy or an additional insulin injection (insulin regime intensification). Among patients situated in the action stage of the Stages of Change Model, Ajzen's Theory of Planned Behavior was used to explore participants' attitudes toward a behavior, their perceived behavioral control, and subjective norms as factors that influence their behavioral intention [37]. Specifically for experiences with and barriers to

the task of insulin titration, Leventhal et al’s Self-Regulatory Theory was used to explore illness representations based on perceptions of cause, identity, timeline, consequences, and controllability of diabetes in terms of prevention and cure [38]. Table 2 gives an overview of all theoretical constructs that were used during data analysis. The two coders (ACRS, WTG) resolved discrepancies through consensus. In case of disagreement, a third researcher (NP) was consulted. Having a third party examine the coding choices made by team members is a technique that has also been used in previous qualitative studies [39]. The entire final codebook is listed in Multimedia Appendix 1.

Two researchers (ACRS and WTG) independently categorized the remarks according to three different topic schemes: (1) the medical task addressed (SMBG, insulin injection, or insulin titration), (2) whether the remark concerned a (perceived) barrier or an experience, and (3) the construct from one of the theories.

When participants reported that an action was not performed, not performed in a timely fashion, or not performed correctly, the reported cause of this circumstance was classified as a barrier to performing the action. We distinguished two types of barriers: barriers that were reported by participants themselves, and barriers that were induced from participants’ remarks by the researchers. To this end, an experienced diabetes physician (ACRS) verified whether patient reported behavior was in line with common recommendations in the field of diabetes, and if this was not the case, the cause of this circumstance was classified as an induced barrier. For example, if patients reported that they did not see advantages of glucose-lowering actions for blood glucose levels that were evidently too high, this was classified as an induced barrier caused by a lack of knowledge of the long-term risks of diabetes or of treatment targets. All patient experiences were classified as “positive” or “negative”, depending on the feelings reported by the participant in question.

Figure 2. Flowchart of matching bottom-up coded remarks to constructs from theoretical frameworks.

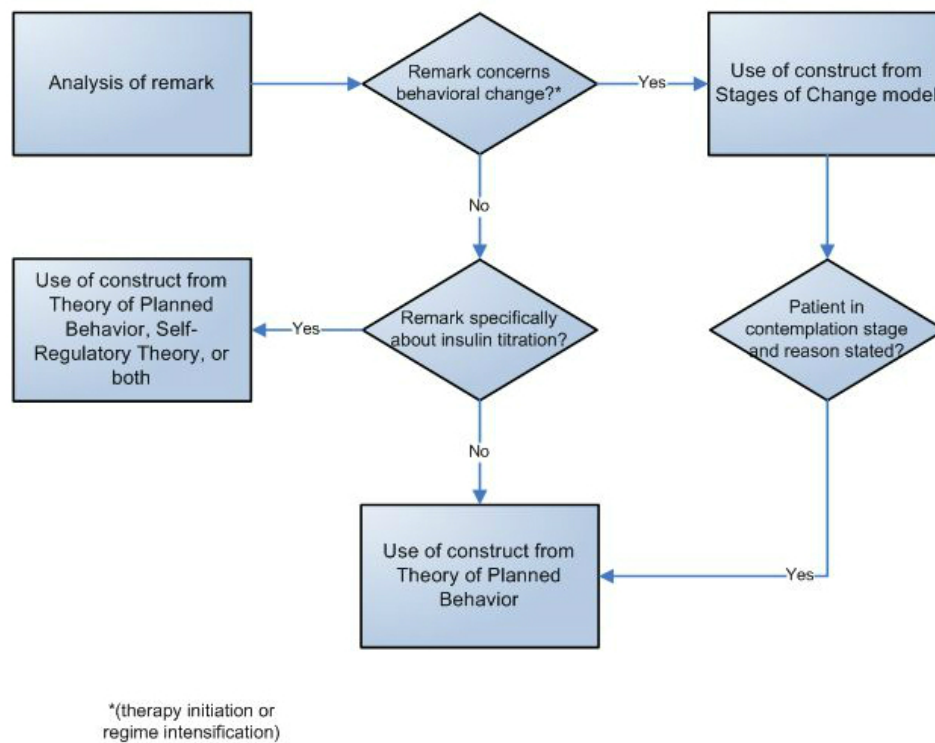


Table 2. Overview of theoretical constructs used in data analysis.

Theory	Focus	Constructs	Description ^a	
Stages of Change Model [36]	Individual's motivation and readiness to change a behavior	Stages	Precontemplation	Not yet acknowledging that there is a problem behavior that needs to be changed
			Contemplation	Acknowledging that there is a problem but not yet ready or sure of wanting to make a change
			Preparation	Getting ready to change
			Action	Changing behavior
			Maintenance	Maintaining the behavior change
Self-Regulatory Theory [38]	Individual's cognitive representations of their current health status based on illness representation	Illness representation	Cause of illness	Perceived causes of an illness
			Identity with the illness	Name or label of an illness
			Consequences	Anticipated and experienced consequences of an illness
			Timeline	Perception of progress and duration of an illness
			Controllability	Perception of amenability to control or cure an illness
Theory of Planned Behavior [37]	Relations between an individual's beliefs, attitudes, intentions, behavior and perceived control over that behavior	Factors	Attitude toward behavior	Positive or negative feelings about performing a behavior
			Subjective norm	Perception of whether people important to the individual think the behavior should be performed
			Perceived behavioral control	Perception of the difficulty of performing a behavior
Integrated model of User Satisfaction and Technology Acceptance [31]	Causal chain from key characteristics of system design to beliefs and expectations about outcomes that ultimately determine usage	Factors	Completeness	Degree to which the system provides all necessary information
			Accuracy	Perception that the information is correct
			Format	Perception of how well the information is presented
			Currency	Perception of the degree to which the information is up to date
			Reliability	Dependability of system operation
			Flexibility	Way the system adapts to changing demands of the patient
			Integration	Way the system allows data to be integrated from various sources
			Accessibility	Ease with which information can be accessed from the system
Timeliness	Degree to which the system offers timely responses to requests for information or action			

^aIn the context of our study, the term “behavior” in this column can be interpreted as either (1) performing self-monitoring of blood glucose, (2) injecting insulin, or (3) titrating insulin.

Analysis of Perceptions of and Experiences With Computer-Assisted Insulin Self-Titration

The perceptions of computer-assisted insulin self-titration that emerged from the face-to-face and telephone interviews were similarly coded bottom-up to eventually develop a codebook. Here too, if consensus could not be reached, a third researcher (NP) was consulted. To support examination of the final coded data of Group 2, two researchers matched each code to a construct from the integrated model of technology acceptance and user satisfaction constructed by Wixom and Todd [31]. This model provides a mechanism for understanding how system and information characteristics influence intended usage. It proposes that system characteristics (reliability, flexibility,

integration, accessibility, and timeliness) and information characteristics (completeness, accuracy, format, and currency) influence system and information quality, respectively. Following a causal chain, system and information quality influence object-based attitudes toward the system that can shape behavioral beliefs of usefulness, ease of use, and, ultimately, system usage. For example, if a computer system is hard to access, this will negatively influence its perceived usefulness and users may decide to use other systems of information sources instead. Table 2 (bottom row) lists the theoretical constructs from this model. The final coded data of Group 1 were not matched to constructs of a model as these patients were never presented with the computer-assisted insulin self-titration system.

Results

Participants

To recruit the 10 participants who had never used a computer-assisted insulin self-titration system, 12 patients were invited. All 10 patients who participated in the pilot study with

PANDIT agreed to take part in this study. Participant characteristics are displayed in Table 3. In the following sections, we first describe the experiences with and barriers to insulin self-management categorized by medical task. Subsequently, we describe patients' perceptions of, and experiences with, computer-assisted insulin self-titration per patient group.

Table 3. Baseline characteristics of study participants by group.

Characteristic	Group 1 (n=10)	Group 2 (n=10)
Demographic characteristics		
Age (years), mean (SD)	53.9 (7.7)	56.9 (8.9)
Male sex, n (%)	6 (60%)	6 (60%)
Education level^a, n (%)		
Low education level	3 (30%)	2 (20%)
Middle education level	4 (40%)	4 (40%)
High education level	3 (30%)	4 (40%)
Diabetes characteristics		
Diabetes duration since diagnosis (years), median (min-max)	8.0 (2.0-25.0)	7.0 (1.0-23.0)
Duration of insulin use (years), median (min-max)	3.0 (0.3-15.0)	1.0 (0.0-3.0)
Insulin regimen		
Basal, n (%)	7 (70%)	10 (100%)
Multiple daily, n (%)	3 (30%)	0 (0%)
HbA1c, mean (SD)	7.2 (0.7)	8.4 (6.3)

^aLow education level: primary school or none; Middle education level: vocational or other secondary school; High education level: university or vocational postsecondary school.

Experiences With and Barriers to Insulin Self-Management

In total, the face-to-face interviews yielded 77 unique remarks regarding experiences with and barriers to insulin self-management (Multimedia Appendix 1). A total of 25 remarks were categorized as barriers, of which six were induced by the diabetes physician. Most barriers concerned the task of insulin titration. For three remarks, no suitable construct was found, for which we added one new construct: "fit between medical commitment and daily routine", implicating the impact of a patient's daily activities on the performance of a medical task. To define patients' attitudes more specifically, we added three child-constructs under the construct "attitude towards behavior" (Theory of Planned Behavior), namely "perceived usefulness", "physical impact", and "effect on daily life". We also added one child-construct under the construct "perceived behavior control" (Theory of Planned Behavior), namely "perceived cognitive load", implicating the emotional or cognitive burden of the medical task. No remarks concerned the constructs "cause of illness" or "identity with the illness" (Self-Regulatory Theory), or "preparation" (Stages of Change Model).

Insulin Therapy Initiation and Insulin Regime Intensification

Five remarks were made about insulin therapy initiation or insulin regime intensification (ie, adding a bolus insulin before a meal). Most patients reported having a reluctance to using an injectable drug due to fear of a lifelong commitment to insulin therapy or fear of disease progression ("contemplation stage", Stages of Change Model).

I just can't stand the injection. If it would be once a day, it would be fine. But I had to inject twice a day. ...Two times is really the maximum for me. They proposed to me to inject three times, but I said no. I don't want that. I couldn't do it. [Female, age 46, Group 1, participant 7]

There were also patients who experienced positively that initiating insulin therapy was effective in lowering blood glucose values.

Self-Monitoring of Blood Glucose

A total of 13 remarks were made about SMBG. Many patients reported that the performance of measurement of blood glucose was accompanied by feelings of pain or tenderness in the fingertips ("physical impact", child construct of "attitude towards behavior", Theory of Planned Behavior). However,

these patients reported that, despite the pain, they still performed the measurements.

I still think finger pricks are a disaster. It just hurts....I have tried different lancets but all of them hurt. ...OK, I have to, I have to, but still I don't like those horrible nasty pricks. [Female, age 60, Group 1, participant 1]

Some patients indicated that the painful feelings were outweighed by the usefulness of SMBG to detect aberrant blood glucose values and to become acquainted with the effects of daily activities and insulin on blood glucose values.

I used to measure 4-6 times a day. At a certain point you get acquainted with hypoglycemia and learn to sense a hypoglycemic event coming up. ...After a year I knew how much I food had to eat and insulin I had to inject to prevent a hypoglycemic event. [Male, age 40, Group 1, participant 4]

However, these patients no longer perceived these frequent SMBG as useful when their blood glucose became better controlled and thus were mostly within the target range.

Insulin Injection

A total of 31 remarks were made about insulin injection. Many patients on insulin therapy mentioned injection pain and site reactions localized on the skin, such as bruises and swellings, as a drawback of insulin therapy (“physical impact”, child construct “attitude towards behavior”, Theory of Planned Behavior). Nevertheless, the physical impact of insulin injections did not necessarily withhold patients from compliant use of insulin therapy.

I inject the rapid-acting insulin in my abdomen and I don't like it. I also have bruises on my abdomen and it's more painful and more sensitive. ...This wouldn't stop me from injecting, but I do avoid the injection site for some days to give it time to recover. [Female, age 59, Group 1, participant 3]

Another negative feeling toward insulin injection that many patients elaborated on was that the commitment to insulin therapy negatively affected their daily personal or social life (“effect on daily life”, child construct “attitude towards behavior”, Theory of Planned Behavior).

If someone asks: do you want to join us? ...I can't, because I did not bring a sandwich ...Or if someone proposes at 5 PM to go to the café and have dinner, I also can't because I did not bring my injection kit. I always have to carefully consider bringing something or not. [Female, age 60, Group 1, participant 1]

Also, several patients mentioned they felt ashamed to inject their insulin in public. Whereas most sought privacy in the restroom to attend their medical needs, others avoided social occasions (“subjective norm”, Theory of Planned Behavior).

If I'm with other people, I dislike injecting. In my personal environment, my wife is the only one who knows [of my disease]. ...I don't feel the necessity to tell others. If I have to inject in public, I retreat. I

have to say I avoid this kind of situation. I don't often go to family parties and such things. [Male, age 40, Group 1, participant 4]

Rapid-acting insulin should be administered just before a meal, except when there is a medical reason (eg, delayed gastric emptying) not to do so. However, some participants reported deviating from this advice. One found it intuitively more logical to inject insulin after the meal (“attitude towards behavior”, Theory of Planned Behavior). Another patient initially injected his insulin before a meal but once experienced a hypoglycemic event when dinner in a restaurant was served later than expected and he had already injected the insulin. Hereafter, he switched to injecting his insulin after the meal. Difficulties with estimating how much one would eat were also named as a reason for deviation (“perceived behavioral control”, Theory of Planned Behavior).

Insulin Titration

A total of 30 remarks were made about insulin titration. In particular, increasing the insulin dose was considered a drawback. Patients said that in the short term, increasing the insulin dose led to a fear of developing hypoglycemia (“consequences”, Self-Regulatory Theory). In the long term, they feared that increasing the insulin dose would lead to disease progression or a transition to a multiple daily insulin regimen (“timeline”, Self-Regulatory Theory).

They said increase [your insulin dose] to 18 [insulin units], increase to 24. Then I thought: it's getting worse. Am I getting sicker? ...I didn't know this was normal in the initial phase. ...It takes some time before you realize how it works. [Male, age 54, Group 1, participant 2]

One patient incorrectly perceived that high doses of insulin might interact with other medication (“perceived behavioral control”, Theory of Planned Behavior). Another patient also indicated being reluctant to increase the insulin dose as this would reduce insight in the remaining function of the pancreas (“controllability”, Self-Regulatory Theory).

If you inject more insulin if your blood sugar is too high, you end up in a situation where you're dependent on [exogenous] insulin. Say you have a steady lifestyle and a fixed insulin dose. ...If your [blood glucose] values decrease, it could mean that the Islets of Langerhans function better. If you constantly switch insulin doses, you can't see if you're doing better. [Male, age 40, Group 1, participant 9]

With regard to self-adjustment of insulin dose, many patients reported that they found it difficult to choose a correct insulin dose. Some patients were not aware of correct glucose targets. Consequently, patients continued their insulin dose when they actually should have increased it, or even decreased their insulin dose when they actually should have continued it.

If my blood glucose value is near 10 or 10.2 [mmol/L], I'm not very alarmed. ...If it would be near 15 [mmol/L], I could call [my diabetes nurse]. ...Until 10 point something it does not bother me. [Male, age 54, Group 1, participant 2]

Also, some patients felt that they were unable to control their blood glucose values due to the large number of external, disturbing factors such as irregular working hours, poor sleeping, and influenza episodes. They would typically consider the evolution of their blood glucose values as unpredictable and varying for no apparent reason (“controllability”, Self-Regulatory Theory).

Other Barriers

Patients indicated that the obligation to use insulin on top of other therapies increased the mental burden of disease management (“perceived cognitive load”, child construct of “perceived behavior control”, Theory of Planned Behavior). Therefore, these patients preferred to limit the frequency of performing any medical procedure to a minimum.

I prefer to perform as few medical procedures as possible. I already take medication because my thyroid was removed, that's why I've become overweight and that's how I got diabetes. Therefore, I try to limit everything to a minimum. Everybody prefers to be healthy and without needs. [Female, age 59, Group 1, participant 3]

Furthermore, daily activities and irregular working hours were named as barriers to integrate SMBG and insulin injection in their daily routine (“fit between medical commitment and daily routine”, added construct). As a result, these tasks were often forgotten or delayed.

Perceptions of and Experiences With Computer-Assisted Insulin Self-Titration

The telephone interviews and face-to-face interviews yielded a total of 20 and 18 remarks respectively regarding perceptions of computer-assisted insulin self-titration. The final coded data of Group 1 were not matched to constructs of a model. These patients mainly discussed whether they found it useful and/or efficient and whether they could trust the system. With regard to Group 2, remarks regarding functionalities that were missing in PANDIT were assigned to an added construct to the Theoretical Integration of User Satisfaction and Technology Acceptance: “missing functionalities”. Furthermore, for three remarks, we added two other constructs: “perceived usefulness”, implicating the degree to which a person believes that his or her diabetes would benefit from the use of a computer-assisted insulin self-titration system; and “attitude towards behavior”, implicating the positive or negative feelings about performing computer-assisted insulin self-titration. The constructs “accuracy”, “currency”, “reliability”, “integration”, and “timeliness” were not applicable.

Patients who had never used a computer-assisted insulin-titration system (Group 1) and who expected a computer-assisted insulin self-titration system to be useful felt they had enough knowledge to know when their insulin should be adjusted, but still believed that the system advice would be useful as it could confirm their reasoning.

When patients were asked how they would feel about receiving insulin dosing advice at home, almost all appreciated that it

could save time and effort in comparison to face-to-face contacts and telephone calls with their caregiver.

It could save me time. Sometimes you have to wait here [in the hospital] for quite a while. ... Other times there are telephone appointments but that also requires waiting if the line is busy. The use of such a system would save both parties time. If it's easy and possible [to titrate the insulin dose] without speaking to the person I would like that, yes. [Male, age 54, Group 1, participant 2]

The main reason for patients who had used PANDIT in a pilot study (Group 2) to find PANDIT useful was that using the system resulted in an evident improvement of their blood glucose levels (“perceived usefulness”, added construct).

Now that I measure my fasting blood glucose values [and enter my fasting blood glucose values in the system], my daily sugar value is around 6 whereas it used to be over 10 [mmol/L]! I'm very happy with the system. ... I only feel better. [Female, age 47, Group 2, participant 11]

Furthermore, some of these patients generally appreciated the ease with which information could be accessed from the system and the fact that professional guidance was offered through the system when their medical situation had changed (“flexibility”).

If I have a pain attack or the flu everything gets disrupted. If I let them [the caregivers] know how the high or low [blood glucose] value occurred, they can intervene if necessary. [Male, age 64, Group 2, participant 18]

Patients of Group 1 who saw no added value of using such a system typically felt confident to perform insulin titration themselves, or were content with current diabetes management practice.

Why would I want to regulate my [basal] insulin? ... I don't think it's useful for me ... perhaps [it would be useful for] people who have difficulties with their diabetes and don't have it for as long as I do. [Female, age 60, Group 1, participant 1]

Some patients of Group 1 expressed their difficulties in trusting insulin dosing advice provided by a computer system, in particular if insulin dosing advice was solely based on blood glucose values. Likewise, patients who had used PANDIT in a pilot study (Group 2) suggested that the system should also take into account other factors that could influence blood glucose level, such as diet and lifestyle (“completeness”).

I think [computer-assisted insulin self-titration] might not be realizable, because the system will not know your exercise rhythm. If I go fishing one day, then I know what insulin dose I'm going to administer. But if I suddenly give up fishing and go for a 90-minute run, I have already entered my data, and they [caregivers] will not know that I'm exercising. [Male, age 56, Group 1, participant 5]

Currently it's [the determination of new insulin doses] only focused on the glucose values, but of course other

influences are nutrition and lifestyle. Would it be useful to extend the system with information on the lifestyle and food intake? [male, age 64, Group 2, participant 18]

Furthermore, patients of Group 2 sometimes missed personal contact when advice was provided.

I miss the personal contact with the diabetes nurse. ...In August, I started with insulin and every 2 weeks I called my diabetes nurse. I had personal contact by telephone, and even came to the hospital a couple of times, and now things go through email. I find contact by telephone more enjoyable than by computer. [female, age 68, Group 2 participant 16]

Discussion

Principal Results

In this study, we identified experiences with and barriers to SMBG, insulin injection, and insulin titration. Furthermore, we explored perceptions of computer-assisted insulin self-titration among patients with T2DM. Some patients who considered starting with insulin therapy feared a lifelong commitment to insulin therapy, disease progression, or had a general reluctance to using an injectable drug. Once on insulin therapy, patients generally experienced it as being effective. Yet, many barriers can arise when performing insulin injection and insulin titration. Insulin injection is painful, a burden to daily life, and some patients considered it embarrassing to inject in public. Also, some patients administered their rapid-acting insulin incorrectly, namely after mealtime instead of before. This behavior was considered more intuitive and provided more control as it is difficult to predict how much one will eat. With regard to insulin titration, most patients found it difficult to increase the insulin dose. This was caused by a perceived increased risk of hypoglycemia, fear of disease progression, and a perceived lack of control, due to the apprehension that increasing the dose might cause problems. Also, some patients did not know how to adjust their insulin, due to ignorance of treatment targets and unpredictable blood glucose levels. Measurement of blood glucose was considered painful, but this did not withhold them from doing it. Furthermore, insulin therapy was found difficult to integrate with other therapies, and difficult to integrate with daily activities or irregular working hours.

Patients who had never used a computer-assisted insulin self-titration system felt they had enough knowledge to know when their insulin should be adjusted, but still believed that they could use the system to verify their reasoning. Furthermore, patients appreciated that receiving insulin dosing advice at their home could save time and effort. Patients who felt confident in performing self-titration found such a system not useful. Patients who had used PANDIT found the system useful when their glycemic regulation improved. Patients appreciated the ease with which information could be accessed from the system. For some patients, the absence of personal contact with their caregiver was a drawback. While guidelines state that adjustment of basal insulin dose based on fasting plasma glucose values is sufficient [17], both patients who had and those who had not

used PANDIT felt that such a system should take more patient data into consideration, such as lifestyle and diet factors.

Comparison With Prior Work

The first studies that reported on patients' experiences with regard to insulin therapy focused on patients with T1DM [40]. However, in the United States, it was estimated that in 1999-2000 almost 30% of people with T2DM were treated with insulin [41]. In line with this increase, recent qualitative studies also focused on patients with T2DM, mainly with regard to the resistance to start insulin therapy [23,42], but also with regard to adherence to insulin therapy [29]. However, these studies were large-scale surveys using self-reported questionnaires with a predefined list of possible reasons for insulin discontinuation [23,29,42-45]. The predefined list was restricted mostly to explore patients' attitudes, either positive or negative, toward insulin therapy and did not cover for example acknowledgment of behavioral change. Furthermore, these studies aimed to collect quantitative information rather than the more "in-depth" and "explorative" qualitative information. The majority of barriers that were found in these studies were also elicited in our study. Important barriers with regard to insulin therapy that were not elicited in our study were: a perceived low efficacy of insulin therapy [23], inadequate explanation of the risks and benefits of insulin therapy by health care provider [42], a physician advising against insulin therapy [45], and the belief that glucose values are under control without insulin therapy [45]. Two previous studies that explored experiences with diabetes self-management of patients with T2DM and also applied in-depth interviews focused on experiences when performing either diet, exercise, or SMBG [24,27]. These studies did not take into regard insulin therapy. Almost all negative experiences referring to SMBG were also elicited in our study. One reason that was not detected in our study causing patients to stop performing SMBG was patients' inference that health professionals did not consider it important [27]. Nowadays, another important focus of study are barriers to diabetes self-management in general among specific patient groups with low adherence rates [46,47]. So far, we do not know of any study that has investigated patients' perceptions with regard to insulin self-titration or computer-assisted insulin self-titration, although patients with T2DM increasingly participate in performing insulin titration [48].

Limitations

A predefined number of 20 patients was included in this study. We did not systematically recruit patients until saturation was achieved, decreasing the likelihood that a complete set of barriers was found. However, the study population consisted of participants with a broad spectrum of age, sex, and HbA1c and was recruited from both general practices and a university hospital. Furthermore, because we had a small number of participants, we were able to perform interviews thoroughly. This increased the likelihood of acquiring a complete set of barriers. Another limitation is the fact that the researchers that conducted the interviews were involved in the development of PANDIT. This may have provoked socially desirable responses by the participants. To increase the reliability of our results, all analyses were separately performed by two researchers. Other

than in previous studies that had used in-depth interviews to elicit obstacles to diabetes self-management [24,27], we used theoretical frameworks for the development of the codebook. This provided us the opportunity to position our findings within the existing body of knowledge. Glaser and Strauss acknowledge that no researcher can totally disregard previous literature or theories [49]. However, using a preconceived model may also have introduced bias or contaminated emerging concepts.

Conclusions

Patients with diabetes on insulin therapy encounter multiple obstacles when performing self-management behavior. Insulin injection itself is painful, a daily recurring burden, and can be embarrassing when performed in the presence of others. Patients experience difficulties with the timing of injections, and with integrating insulin therapy with other therapies and daily activities. Adjustment of insulin dose is a challenge; some patients are ignorant of treatment targets and find it difficult to cope with unpredictable blood glucose levels. In particular, patients have difficulties increasing the insulin dose because they fear hypoglycemia and because they associate higher insulin doses with disease progression. Computer-assisted self-titration can increase patient awareness of treatment targets and increase their confidence in self-adjusting the insulin dose. Nevertheless, some barriers may persist when using computer-assisted insulin self-titration systems, such as reluctance to increase the insulin dose. Furthermore, computer-assisted self-titration could also introduce new barriers. There might be a lack of trust in a computerized insulin dosing advice that, based on guidelines, is primarily based on blood glucose values and does not take account of other patient data, such as lifestyle and diet. Also, for some patients, the absence of personal contact with their caregiver when advice is provided can be a drawback.

Recommendations for Care Practice

For the design and future implementation of computer-assisted insulin titration systems for patients, we have four recommendations. First, in order to increase the effectiveness of the system in lowering glucose values, the caregiver should verify if the patient is both willing and able to perform SMBG, insulin injections, and to log into a computer-assisted system for insulin dosing advice. Second, to increase acceptance of dosing advice by the system, patients should have the possibility to increase professional involvement through the system, such as by telemedicine functionalities that allow easily accessible consultations with the caregiver. Third, to minimize the burden on a patient's daily life, the frequency of consulting the system and blood glucose measurements should be decreased once patients have reached near-normal glucose levels. And finally, patients should be motivated to start and continue use of the computer-assisted system by emphasizing its main advantage: achieving lower blood glucose values. This should be emphasized by both the system, for example, by means of a graphical display of blood sugar values, and the caregiver.

Recommendations for Future Research

We would recommend future research to explore the wishes and needs of patients to interact with their caregiver during the process of insulin titration, in particular, what situations trigger their need to consult the caregiver and what communication modalities would be preferred. Following these potential additional features to the computer-assisted insulin titration system, such as telemedicine functionalities to increase patient-provider interaction, future research should investigate how this would affect patient acceptance.

Acknowledgments

ACRS and WTG contributed to the design of the study, collected data, analyzed and interpreted data, and drafted the manuscript. NP contributed to the design of the study, the analysis and interpretation of the data, and reviewed/edited the manuscript. FH and JBLH reviewed/edited the manuscript. All authors significantly contributed to the work and approved the manuscript for submission. ACRS and NP are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

The present study was financially supported by the Netherlands Organization for Health Research and Development (ZonMw). The study sponsor had no role in the design and conduct of the study, collection, management, analysis, and interpretation of the data, nor in preparation, review, and approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Codebook.

[PDF File (Adobe PDF File), 92KB - [jmir_v16i10e235_app1.pdf](#)]

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004 May;27(5):1047-1053. [Medline: [15111519](#)]

2. Rosenstock J, Davies M, Home PD, Larsen J, Koenen C, Schernthaner G. A randomised, 52-week, treat-to-target trial comparing insulin detemir with insulin glargine when administered as add-on to glucose-lowering drugs in insulin-naïve people with type 2 diabetes. *Diabetologia* 2008 Mar;51(3):408-416 [FREE Full text] [doi: [10.1007/s00125-007-0911-x](https://doi.org/10.1007/s00125-007-0911-x)] [Medline: [18204830](https://pubmed.ncbi.nlm.nih.gov/18204830/)]
3. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 1998 Sep 12;352(9131):837-853. [Medline: [9742976](https://pubmed.ncbi.nlm.nih.gov/9742976/)]
4. Nathan DM, Buse JB, Davidson MB, Ferrannini E, Holman RR, Sherwin R, et al. Management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: update regarding thiazolidinediones: a consensus statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2008 Jan;31(1):173-175. [doi: [10.2337/dc08-9016](https://doi.org/10.2337/dc08-9016)] [Medline: [18165348](https://pubmed.ncbi.nlm.nih.gov/18165348/)]
5. Gulliford MC, Ashworth M, Robotham D, Mohiddin A. Achievement of metabolic targets for diabetes by English primary care practices under a new system of incentives. *Diabet Med* 2007 May;24(5):505-511. [doi: [10.1111/j.1464-5491.2007.02090.x](https://doi.org/10.1111/j.1464-5491.2007.02090.x)] [Medline: [17381507](https://pubmed.ncbi.nlm.nih.gov/17381507/)]
6. Saaddine JB, Cadwell B, Gregg EW, Engelgau MM, Vinicor F, Imperatore G, et al. Improvements in diabetes processes of care and intermediate outcomes: United States, 1988-2002. *Ann Intern Med* 2006 Apr 4;144(7):465-474. [Medline: [16585660](https://pubmed.ncbi.nlm.nih.gov/16585660/)]
7. Inspectie voor de Gezondheidszorg (IGZ). Prestatie-indicatoren als onafhankelijke graadmeter voor de kwaliteit van in ziekenhuizen verleende zorg. Het resultaat telt 2006. Den Haag: IGZ; 2007. URL: <http://www.igz.nl/actueel/nieuws/hetresultaatteletprestatiesvanziekenhuizenin2006.aspx> [accessed 2014-09-28] [WebCite Cache ID 6Svit4abq]
8. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, American Diabetes Association (ADA), European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care* 2012 Jun;35(6):1364-1379 [FREE Full text] [doi: [10.2337/dc12-0413](https://doi.org/10.2337/dc12-0413)] [Medline: [22517736](https://pubmed.ncbi.nlm.nih.gov/22517736/)]
9. Goldstein DE, Little RR, Lorenz RA, Malone JJ, Nathan DM, Peterson CM, American Diabetes Association. Tests of glycemia in diabetes. *Diabetes Care* 2003 Jan;26 Suppl 1:S106-S108. [Medline: [12502632](https://pubmed.ncbi.nlm.nih.gov/12502632/)]
10. Strange P. Treat-to-target insulin titration algorithms when initiating long or intermediate acting insulin in type 2 diabetes. *J Diabetes Sci Technol* 2007 Jul;1(4):540-548 [FREE Full text] [Medline: [19885117](https://pubmed.ncbi.nlm.nih.gov/19885117/)]
11. Harris MI, National Health and Nutrition Examination Survey (NHANES III). Frequency of blood glucose monitoring in relation to glycemic control in patients with type 2 diabetes. *Diabetes Care* 2001 Jun;24(6):979-982. [Medline: [11375356](https://pubmed.ncbi.nlm.nih.gov/11375356/)]
12. Cramer JA. A systematic review of adherence with medications for diabetes. *Diabetes Care* 2004 May;27(5):1218-1224. [Medline: [15111553](https://pubmed.ncbi.nlm.nih.gov/15111553/)]
13. 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](https://pubmed.ncbi.nlm.nih.gov/12435261/)]
14. Kawamura T. The importance of carbohydrate counting in the treatment of children with diabetes. *Pediatr Diabetes* 2007 Oct;8 Suppl 6:57-62. [doi: [10.1111/j.1399-5448.2007.00287.x](https://doi.org/10.1111/j.1399-5448.2007.00287.x)] [Medline: [17727386](https://pubmed.ncbi.nlm.nih.gov/17727386/)]
15. Brown LL, Lustria ML, Rankins J. A review of web-assisted interventions for diabetes management: maximizing the potential for improving health outcomes. *J Diabetes Sci Technol* 2007 Nov;1(6):892-902 [FREE Full text] [Medline: [19885163](https://pubmed.ncbi.nlm.nih.gov/19885163/)]
16. Schmidt S, Meldgaard M, Serifovski N, Storm C, Christensen TM, Gade-Rasmussen B, et al. Use of an automated bolus calculator in MDI-treated type 1 diabetes: the BolusCal Study, a randomized controlled pilot study. *Diabetes Care* 2012 May;35(5):984-990 [FREE Full text] [doi: [10.2337/dc11-2044](https://doi.org/10.2337/dc11-2044)] [Medline: [22344610](https://pubmed.ncbi.nlm.nih.gov/22344610/)]
17. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, American Diabetes Association (ADA), European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). *Diabetes Care* 2012 Jun;35(6):1364-1379 [FREE Full text] [doi: [10.2337/dc12-0413](https://doi.org/10.2337/dc12-0413)] [Medline: [22517736](https://pubmed.ncbi.nlm.nih.gov/22517736/)]
18. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2013;3:CD008776. [doi: [10.1002/14651858.CD008776.pub2](https://doi.org/10.1002/14651858.CD008776.pub2)] [Medline: [23543567](https://pubmed.ncbi.nlm.nih.gov/23543567/)]
19. Ramadas A, Quek KF, Chan CK, Oldenburg B. Web-based interventions for the management of type 2 diabetes mellitus: a systematic review of recent evidence. *Int J Med Inform* 2011 Jun;80(6):389-405. [doi: [10.1016/j.ijmedinf.2011.02.002](https://doi.org/10.1016/j.ijmedinf.2011.02.002)] [Medline: [21481632](https://pubmed.ncbi.nlm.nih.gov/21481632/)]
20. Lim S, Kang SM, Shin H, Lee HJ, Won Yoon J, Yu SH, et al. Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. *Diabetes Care* 2011 Feb;34(2):308-313 [FREE Full text] [doi: [10.2337/dc10-1447](https://doi.org/10.2337/dc10-1447)] [Medline: [21270188](https://pubmed.ncbi.nlm.nih.gov/21270188/)]
21. Korytkowski M. When oral agents fail: practical barriers to starting insulin. *Int J Obes Relat Metab Disord* 2002 Sep;26 Suppl 3:S18-S24. [doi: [10.1038/sj.ijo.0802173](https://doi.org/10.1038/sj.ijo.0802173)] [Medline: [12174319](https://pubmed.ncbi.nlm.nih.gov/12174319/)]
22. Kunt T, Snoek FJ. Barriers to insulin initiation and intensification and how to overcome them. *Int J Clin Pract Suppl* 2009 Oct(164):6-10. [doi: [10.1111/j.1742-1241.2009.02176.x](https://doi.org/10.1111/j.1742-1241.2009.02176.x)] [Medline: [19751453](https://pubmed.ncbi.nlm.nih.gov/19751453/)]

23. Peyrot M, Rubin RR, Lauritzen T, Skovlund SE, Snoek FJ, Matthews DR, International DAWN Advisory Panel. Resistance to insulin therapy among patients and providers: results of the cross-national Diabetes Attitudes, Wishes, and Needs (DAWN) study. *Diabetes Care* 2005 Nov;28(11):2673-2679. [Medline: [16249538](#)]
24. Hall RF, Joseph DH, Schwartz-Barcott D. Overcoming obstacles to behavior change in diabetes self-management. *Diabetes Educ* 2003;29(2):303-311. [Medline: [12728757](#)]
25. Karter AJ, Ferrara A, Darbinian JA, Ackerson LM, Selby JV. Self-monitoring of blood glucose: language and financial barriers in a managed care population with diabetes. *Diabetes Care* 2000 Apr;23(4):477-483 [FREE Full text] [Medline: [10857938](#)]
26. Wagner J, Malchoff C, Abbott G. Invasiveness as a barrier to self-monitoring of blood glucose in diabetes. *Diabetes Technol Ther* 2005 Aug;7(4):612-619. [doi: [10.1089/dia.2005.7.612](#)] [Medline: [16120035](#)]
27. Peel E, Douglas M, Lawton J. Self monitoring of blood glucose in type 2 diabetes: longitudinal qualitative study of patients' perspectives. *BMJ* 2007 Sep 8;335(7618):493 [FREE Full text] [doi: [10.1136/bmj.39302.444572.DE](#)] [Medline: [17761996](#)]
28. Rubin RR, Peyrot M, Kruger DF, Travis LB. Barriers to insulin injection therapy: patient and health care provider perspectives. *Diabetes Educ* 2009;35(6):1014-1022. [doi: [10.1177/0145721709345773](#)] [Medline: [19934459](#)]
29. Davies MJ, Gagliardino JJ, Gray LJ, Khunti K, Mohan V, Hughes R. Real-world factors affecting adherence to insulin therapy in patients with Type 1 or Type 2 diabetes mellitus: a systematic review. *Diabet Med* 2013 May;30(5):512-524. [doi: [10.1111/dme.12128](#)] [Medline: [23323988](#)]
30. Simon AC, Holleman F, Gude WT, Hoekstra JB, Peute LW, Jaspers MW, et al. Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes mellitus. *Artif Intell Med* 2013 Sep;59(1):23-31. [doi: [10.1016/j.artmed.2013.04.009](#)] [Medline: [23735522](#)]
31. Wixom BH, Todd PA. A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research* 2005 Mar;16(1):85-102. [doi: [10.1287/isre.1050.0042](#)]
32. Riddle MC, Rosenstock J, Gerich J, Insulin Glargine 4002 Study Investigators. The treat-to-target trial: randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients. *Diabetes Care* 2003 Nov;26(11):3080-3086. [Medline: [14578243](#)]
33. VERBI GmbH. MAXQDA qualitative data analysis software, 1989-2014. Berlin, Germany; 2014. URL: <http://www.maxqda.com/> [accessed 2014-09-28] [WebCite Cache ID 6SvkVuB2I]
34. Green J, Thorogood N. Qualitative methods for health research (Introducing Qualitative Methods series). London: Sage Publications Ltd; 2004.
35. DeCuir-Gunby JT, Marshall PL, McCulloch AW. Developing and using a codebook for the analysis of interview data: an example from a professional development research project. *Field Methods* 2010 Dec 27;23(2):136-155. [doi: [10.1177/1525822X10388468](#)]
36. 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](#)]
37. Ajzen I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 1991 Dec;50(2):179-211. [doi: [10.1016/0749-5978\(91\)90020-T](#)]
38. Leventhal H, Leventhal EA, Contrada RJ. Self-regulation, health, and behavior: a perceptual-cognitive approach. *Psychology & Health* 1998 Jul;13(4):717-733. [doi: [10.1080/08870449808407425](#)]
39. Roebuck A, Furze G, Thompson DR. Health-related quality of life after myocardial infarction: an interview study. *J Adv Nurs* 2001 Jun;34(6):787-794. [Medline: [11422549](#)]
40. Ary DV, Toobert D, Wilson W, Glasgow RE. Patient perspective on factors contributing to nonadherence to diabetes regimen. *Diabetes Care* 1986;9(2):168-172. [Medline: [3698782](#)]
41. Koro CE, Bowlin SJ, Bourgeois N, Fedder DO. Glycemic control from 1988 to 2000 among U.S. adults diagnosed with type 2 diabetes: a preliminary report. *Diabetes Care* 2004 Jan;27(1):17-20. [Medline: [14693960](#)]
42. Karter AJ, Subramanian U, Saha C, Crosson JC, Parker MM, Swain BE, et al. Barriers to insulin initiation: the translating research into action for diabetes insulin starts project. *Diabetes Care* 2010 Apr;33(4):733-735 [FREE Full text] [doi: [10.2337/dc09-1184](#)] [Medline: [20086256](#)]
43. Broadbent E, Donkin L, Stroh JC. Illness and treatment perceptions are associated with adherence to medications, diet, and exercise in diabetic patients. *Diabetes Care* 2011 Feb;34(2):338-340 [FREE Full text] [doi: [10.2337/dc10-1779](#)] [Medline: [21270191](#)]
44. Peyrot M, Barnett AH, Meneghini LF, Schumm-Draeger PM. Insulin adherence behaviours and barriers in the multinational Global Attitudes of Patients and Physicians in Insulin Therapy study. *Diabet Med* 2012 May;29(5):682-689 [FREE Full text] [doi: [10.1111/j.1464-5491.2012.03605.x](#)] [Medline: [22313123](#)]
45. Oliveria SA, Menditto LA, Ulcickas Yood M, Koo YH, Wells KE, McCarthy BD. Barriers to the initiation of, and persistence with, insulin therapy. *Curr Med Res Opin* 2007 Dec;23(12):3105-3112. [doi: [10.1185/030079907X242638](#)] [Medline: [17977482](#)]
46. Hu J, Amirehsani K, Wallace DC, Letvak S. Perceptions of barriers in managing diabetes: perspectives of Hispanic immigrant patients and family members. *Diabetes Educ* 2013;39(4):494-503 [FREE Full text] [doi: [10.1177/0145721713486200](#)] [Medline: [23640301](#)]

47. Onwudiwe NC, Mullins CD, Winston RA, Shaya FT, Pradel FG, Laird A, et al. Barriers to self-management of diabetes: a qualitative study among low-income minority diabetics. *Ethn Dis* 2011;21(1):27-32. [Medline: [21462726](#)]
48. Davies M, Storms F, Shutler S, Bianchi-Biscay M, Gomis R, ATLANTUS Study Group. Improvement of glycemic control in subjects with poorly controlled type 2 diabetes: comparison of two treatment algorithms using insulin glargine. *Diabetes Care* 2005 Jun;28(6):1282-1288. [Medline: [15920040](#)]
49. Glaser BG, Strauss AL. *The discovery of grounded theory; strategies for qualitative research*. Chicago, IL: Aldine Pub. Co; 2012.

Abbreviations

FPG: fasting plasma glucose

HbA1c: glycated hemoglobin

PANDIT: Patient Assisting Net-based Diabetes Insulin Titration system

SMBG: self-monitoring of blood glucose

T1DM: type 1 diabetes mellitus

T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 24.12.13; peer-reviewed by M Kazmer, G Burnett; comments to author 28.04.14; revised version received 24.06.14; accepted 10.07.14; published 23.10.14.

Please cite as:

Simon ACR, Gude WT, Holleman F, Hoekstra JBL, Peek N

Diabetes Patients' Experiences With the Implementation of Insulin Therapy and Their Perceptions of Computer-Assisted Self-Management Systems for Insulin Therapy

J Med Internet Res 2014;16(10):e235

URL: <http://www.jmir.org/2014/10/e235/>

doi: [10.2196/jmir.3198](https://doi.org/10.2196/jmir.3198)

PMID: [25340869](https://pubmed.ncbi.nlm.nih.gov/25340869/)

©Airin CR Simon, Wouter T Gude, Frits Holleman, Joost BL Hoekstra, Niels Peek. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 23.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evaluating the Interactive Web-Based Program, Activate Your Heart, for Cardiac Rehabilitation Patients: A Pilot Study

Christopher Brough¹, BSc (Hons); Sally Boyce¹, MSc; Linzy Houchen-Wolloff¹, PhD; Louise Sewell¹, PhD; Sally Singh¹, PhD

Centre for Exercise and Rehabilitation Science, University Hospitals of Leicester NHS Trust, Leicester, United Kingdom

Corresponding Author:

Christopher Brough, BSc (Hons)
Centre for Exercise and Rehabilitation Science
University Hospitals of Leicester NHS Trust
Glenfield Hospital
Groby Road
Leicester, LE3 9QP
United Kingdom
Phone: 44 01162502760
Fax: 44 01162502743
Email: christopher.brough@uhl-tr.nhs.uk

Abstract

Background: Conventional cardiac rehabilitation (CR) programs are traditionally based on time-constrained, structured, group-based programs, usually set in hospitals or leisure centers. Uptake for CR remains poor, despite the ongoing evidence demonstrating its benefits. Additional alternative forms of CR are needed. An Internet-based approach may offer an alternative mode of delivering CR that may improve overall uptake. Activate Your Heart (AYH) is a Web-based CR program that has been designed to support individuals with coronary heart disease (CHD).

Objective: The aim of this pilot study was to observe the outcome for participants following the AYH program.

Methods: We conducted a prospective observational trial, recruiting low-risk patients with CHD. Measures of exercise, exercise capacity, using the Incremental Shuttle Walk Test (ISWT), dietary habits, and psychosocial well-being were conducted by a CR specialist at baseline and at 8 weeks following the Web-based intervention.

Results: We recruited 41 participants; 33 completed the program. We documented significant improvements in the ISWT distance (mean change 49.69 meters, SD 68.8, $P < .001$), and Quality of Life (QOL) (mean change 0.28, SD 0.4, $P < .001$). Dietary habits improved with an increased proportion of patients consuming at least 5 portions of fruit and vegetables per day, (22 [71%] to 29 [94%] $P = .01$) and an increased proportion of patients consuming at least 2 portions of oily fish per week (14 [45%] to 21 [68%], $P = .01$). We did not detect changes in anxiety and depression scores or exercise behavior.

Conclusions: We observed important improvements in exercise capacity, QOL, and dietary habits in a group of participants following a Web-based CR program. The program may offer an alternative approach to CR. A mobile version has been developed and we need to conduct further trials to establish its value compared to supervised CR.

(*J Med Internet Res* 2014;16(10):e242) doi:[10.2196/jmir.3027](https://doi.org/10.2196/jmir.3027)

KEYWORDS

Activate Your Heart; coronary heart disease; cardiac rehabilitation; Internet; Web-based

Introduction

Coronary heart disease (CHD) is the world's leading cause of mortality [1]. Cardiac rehabilitation (CR) has been shown to reduce cardiac mortality by 26% and the incidence of further cardiac events [2]. The World Health Organization defines CR as "activities that favorably influence the underlying cause of

the disease and provision of the best possible, physical, mental, and social conditions, so that patients may, by their own efforts, resume as normal a place as possible in the community" [3]. It is characterized by a package of exercise and multidisciplinary education and secondary prevention advice. However, uptake for CR remains poor in many countries [4], despite national targets for uptake. Conventional CR programs are based on time

constrained, structured, group-based programs, usually set in hospitals or local leisure centers. A number of barriers for attending CR have been identified, such as distance to travel, lack of transport [5], patients' reluctance to participate in a group environment, and time constraints due to work commitments [6,7]. Service factors such as lack of capacity, a failure to offer genuine choice of rehabilitation options [8], and poor adoption of technology within CR [9] have also been highlighted as factors contributing to poor uptake.

The limited uptake of conventional CR would suggest that alternative forms of delivering CR need to be developed [10]. Other forms of CR, such as home-based CR, have been tested against conventional CR programs, with the same success [11-13]. These programs offer many of the elements of conventional CR; they are able to provide a program of prescribed exercise and interdisciplinary education, usually as a manual. However, only a very small percentage of those eligible for CR are offered these programs.

There is considerable interest in digital health as a means of delivering care for individuals with long-term conditions of which CHD is an important one, delivering a standard intervention that is not geographically or time restrained. An increasing proportion of retired people are using the Internet [14], reflecting the typical rehabilitation population. In the United Kingdom, 71.9% and 63.7% of males and females aged between 65-74 years respectively have used the Internet [15]. Interestingly, evidence suggests that the number of people seeking health care guidance via the Internet is increasing [16,17].

The use of the Internet allows for greater flexibility as patients are able to complete their CR program at a place and time that suits them. It is also capable of reaching a wider audience, especially those patients that live in rural areas [18]. This growth in Internet use makes it worthy of consideration for intervention development. Previous research has identified the benefits of using Web-based interventions for areas such as weight loss, smoking, physical activity, and reducing depression and anxiety in long-term conditions such as diabetes and asthma [19-23]. Studies have highlighted how Web-based interventions can also help to improve knowledge for patients with chronic health conditions [24].

There is little literature describing initiatives that are relevant to patients with CHD. Studies carried out across Europe, the United States, and Canada have investigated the efficacy of Web-based interventions for those with heart disease. The largest was reported recently in the British Medical Journal [25]; however, it was not a comprehensive rehabilitation program and recruited participants with a broad range of cardio-vascular diseases and was largely inconclusive. An interesting paper [26] reported on a study from Canada that recruited exclusively people post primary percutaneous coronary revascularization (PCI or angioplasty) who were offered a physical activity intervention that was Web-based. The paper reported a benefit in the intervention arm not observed in the control arm, suggesting the potential value of Web-based interventions in this population. An earlier study [27] randomized patients to an Internet-based case management trial providing risk factor

management. The intervention group had an increase in weight loss compared to the control group, but no additional benefits were observed in depression scores, minutes of exercise, and dietary habits, the outcomes traditionally associated with cardiac rehabilitation. Another small pilot study (n=15) in Canada compared an online comprehensive CR program with standard CR [28], which showed positive benefits for patients by increasing exercise levels and modifying lifestyle behaviors.

There may also be benefits to the service, releasing capacity for CR specialists to manage more complex patients in conventional CR, as well as providing additional choice for those unwilling to do standard CR [27].

Over the last few years, we have developed and refined an online comprehensive CR program, Activate Your Heart (AYH). The aim of this study was to observe the impact on exercise performance and health-related quality of life on patients who completed AYH.

Method

Overview

This was a prospective observational study. Patients were recruited from those referred to CR at the University Hospitals of Leicester NHS Trust. Assessments by a CR specialist were conducted at baseline and after 8 weeks of following AYH. We recruited patients with CHD having one or more of the following: percutaneous coronary intervention (PCI) in the last 3 months, undergone coronary artery bypass graft (CABG) in the last 3 months, or were medically managed for their CHD. To identify low-risk patients, a prospective threshold of at least level 7 (420 meters) of the Incremental Shuttle Walking Test (ISWT) [29] was agreed. Exclusion criteria were unstable angina, significant anxiety/depression ≥ 11 on Hospital Anxiety Depression Scale (HADS) [30], moderate-severe left ventricular dysfunction, inability to perform physical activity due to significant comorbidities, such as severe arthritis, neurological disorders, psychiatric disorders, and not being computer literate.

Intervention

AYH (Figure 1) is an interactive Web-based CR program that offers comprehensive secondary prevention education together with access to CR specialists via a private messaging facility. It was developed by CR specialists and patients who shaped and informed the final product, at the University Hospitals of Leicester NHS Trust. The program is password protected; each participant was given their own unique password to access the AYH program. All participants were able to record and monitor their exercises, and participate with interactive, secondary prevention advice to promote healthy lifestyle changes and reduce risk factors for CHD, via the website (Figure 2).

The AYH program had been structured so as to guide the user through four stages that each have specific tasks the user needs to achieve before progressing to the next stage (Figure 3). Tasks included creating and updating their own short-term goals, completing knowledge tests on CHD and the risk factors, and reading specific topics such as goal setting and making lifestyle changes. The educational reading material includes videos and covers topics such as anatomy and physiology of the heart,

CHD, risk factors for CHD, cardiac tests, and treatments for CHD (Figure 2).

In Stage 1, participants were asked to do a multiple choice questionnaire to establish their knowledge regarding the principles of exercising safely. A score of 80% was set as a threshold to ensure understanding of these principles. In Stages 2-4, participants were required to record all their exercises in an exercise diary. During Stage 2, participants were advised to record a cumulative total of 30 minutes of exercise 5 days out of 7. The intensity of the exercise was based on their performance on the baseline exercise tests and prescribed at 60-80% of baseline performance. In Stage 3, this was amended to 30 minutes of continuous exercise. Finally, in Stage 4, participants were required to record at least 30 minutes of at least moderate exercise 5 days out of 7, in order to fulfill the national requirements for exercise. There is also interactivity around diet and weight management, stress management, and smoking cessation, if appropriate. For smokers, a cost calculator was developed that would calculate how much the user had spent or saved since starting the program. This was delivered along with advice and support to stop smoking.

Other features embedded within the AYH program included a forum where patients were also able to share views and experiences with other program users, a blog, and a frequently asked questions section. The forum was monitored and moderated, as necessary, by CR specialists. In addition patients were able to communicate privately with a CR specialist via the Ask the Expert messaging facility. All queries were answered by a CR specialist within 48 hours of being posted.

The CR team members were provided with individual passwords to access the administration section of the program; this allowed them to view and monitor individual patients' progress and view patient login data. If a patient had not logged in for more than 7 days, or were logging in but not progressing, they would be contacted via email or telephone by the CR specialists. All participants attended a follow-up assessment on completion of the AYH program by a CR specialist.

All data captured on the program were encrypted to safeguard patient confidentiality.

Figure 1. Activate Your Heart homepage.

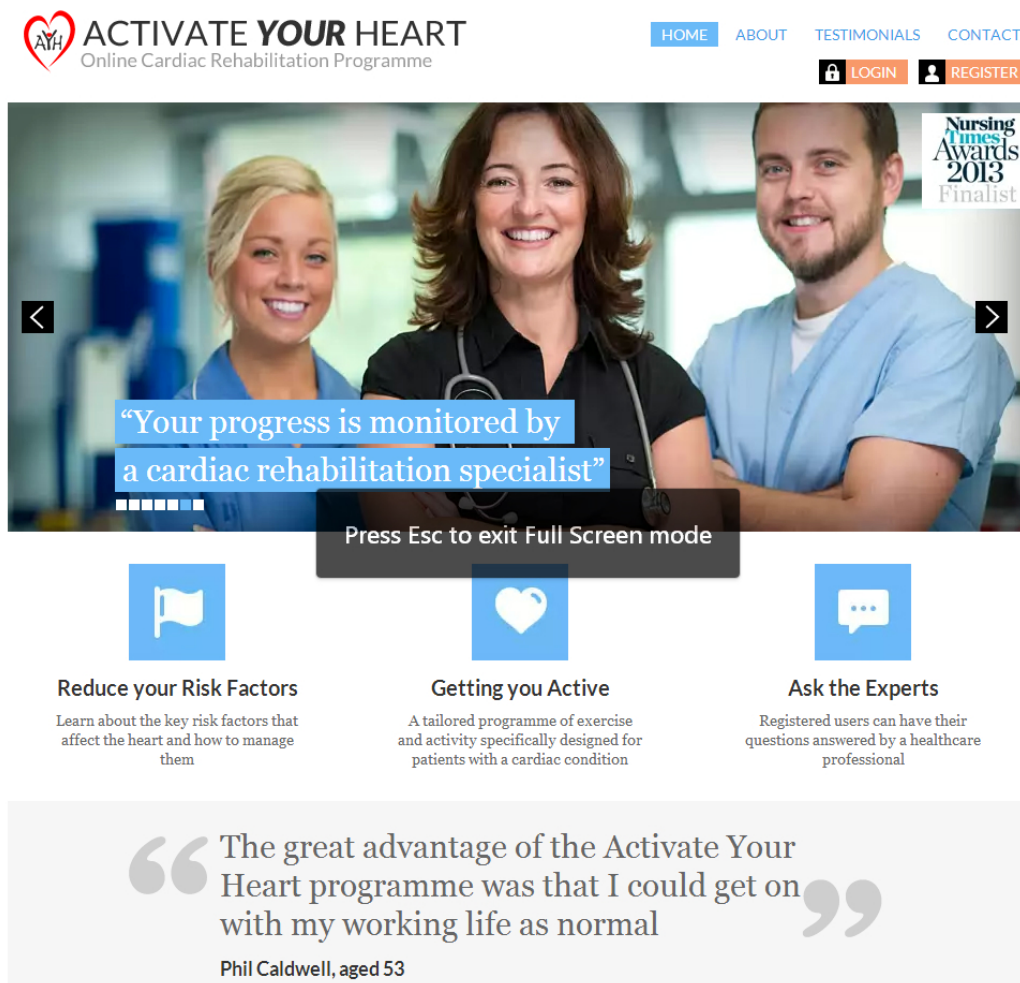


Figure 2. Reading material section.

Reading Material

> All information regarding the heart, risk factors, cardiac tests, treatments and conditions can be found here.




Index
Stage One
Stage Two
Stage Three
Stage Four
Risk Factors
Medical Knowledge
Glossary

Last page read:
[Patient AYH testimonials](#) (Stage One > Patient AYH testimonials)

Watch the video

STAGE ONE




This walking programme aims to help you improve your stamina. It is designed to help you gradually increase the duration (length) of your exercise programme. To achieve the health benefits, it is recommended that you exercise continuously at a moderate pace for 30 minutes a day.

MORE TOPICS

- [Making changes](#)
- [Goal Setting](#)
- [Exercise - Getting started](#)
- [Patient video testimonials](#)
- [Patient AYH testimonials](#)

RISK FACTORS




It is important that you and those around you understand why you have heart disease. Health-care professionals and the media often describe your risk of further heart disease by 'risk factors'. It is important to change your risk factors to reduce the risk of further changes to your heart.

MORE TOPICS

- [Introduction](#)
- [Age](#)
- [Eating for a healthy heart](#)
- [Alcohol](#)
- [Cholesterol](#)
- [Diabetes and heart disease](#)
- [Family history](#)
- [High blood pressure](#)
- [Lack of exercise](#)
- [Smoking](#)
- [South Asians and heart disease](#)
- [Stress](#)
- [Women and heart disease](#)
- [Overweight](#)

MEDICAL KNOWLEDGE



Chest pain or discomfort is often due to conditions that are not related to heart problems (for example, chest infections, injury, or inflammation). However, the chest pain or discomfort associated with heart disease, particularly angina or myocardial infarction (heart attack), has several typical characteristics, which should prompt you to get medical attention.

MORE TOPICS

- [Chest pain: The Facts](#)
- [How the heart works](#)
- [Heart Conditions](#)
- [Cardiac Tests](#)
- [Cardiac Treatments](#)

Figure 3. Part of the dashboard showing what tasks need to be completed in stage 2.

ACTIVATE YOUR HEART

Dashboard Reading Material Forum Blog FAQs Ask the Expert Search GO

Welcome back Chris
You are currently on Week 3 of your programme

Chris B [Logout]
Current Stage: 1 2 3 4
(since 16/10/2012)

STAGE TWO

1 Task: Use Exercise Diary ✓

2 Task: Update your Goals ✓

3 Task: 30 minutes (cumulative) exercise a day over 5 days

4 Test: Complete Quiz 🔒

5 Task: Lifestyle Questions 🔒

CLOSE

Outcomes

Overview

The outcome measures employed were the Hospital Anxiety and Depression Scale (HADS), the MacNew Quality of Life (QOL) questionnaire [31], the ISWT, smoking status, diet (consumption of fruit and vegetables/day and oily fish/week), weight, Body Mass Index (BMI), and number of patients exercising for 30 minutes, 5 days per week. Baseline measurements were taken 4-6 weeks post discharge from hospital. These measurements were then repeated following completion of the AYH program.

Hospital Anxiety and Depression Scale

The HADS is a validated measure of anxiety and depression [32]. It consists of 14 statements, 7 for anxiety and 7 for depression. Each statement is rated on a Likert scale ranging from 0-3. A total score of 8-10 is classed as mild depression or anxiety, 11-14 as moderate depression or anxiety, and ≥ 15 severe depression or anxiety.

MacNew Quality of Life Questionnaire

The MacNew comprises 27 questions, which are divided into 3 domains: emotional, physical, and social. Each question is based on a score between 1 and 7, with lower scores corresponding to impaired QOL. This questionnaire has been reported to be both a valid and reliable measure for this population [33].

Incremental Shuttle Walking Test

The ISWT is a standardized field exercise test, which has shown to be both a valid and reliable test for assessing exercise capacity

and has been previously reported in trials of CR [34,35]. Participants are required to walk up and down a 10-meter course at a pace increasing at 1-minute intervals. The test consists of 12 levels, with 1020 meters as the maximum that can be achieved. All participants were monitored using cardiac portable telemetry during the test. Healthy reference values have been recently described [36].

Statistical Analysis

All data were analyzed using PASW Statistics for Windows, Version 18.0. For normally distributed data, parametric tests were carried out. Within-group changes after the AYH program in HADS scores, ISWT distance and MacNew scores, BMI and weight were analyzed using a paired *t* test. Changes in proportions after using the AYH program (eg, percentage of patients exercising for at least 30 minutes, 5 times per week; percentage of patients eating at least 2 portions of oily fish per week; percentage of patients eating at least 5 portions fruit and vegetables per day; and smoking status) were assessed using the non-parametric, paired samples McNemar's test, as the data were categorical/binary. A *P* value of $P \leq .05$ was reported as significant. Effect size was calculated as a mean change (pre-post) divided by the standard deviation of the change. The magnitude of change can be assessed against the following criteria: small, 0.2 to 0.5; moderate, 0.5 to 0.8; and large, >0.8 [37]. No formal power calculations were used as this was an observational study to assess uptake and collect data on clinical effectiveness.

Ethics

The study was approved by Leicestershire, Northamptonshire and Rutland Research Ethics Committee 2 (approval No.

07/Q2501/114) (ID No. UHL 10322). All subjects for the study provided written informed consent.

Results

We recruited 41 patients to the AYH program (Table 1). Patients presented with low levels of anxiety and depression and higher exercise capacity than is routinely observed in our conventional CR. Of the 41 patients who started the AYH program, 33 attended the follow-up assessment (Tables 2 and 3). There were statistically significant improvements in the ISWT (mean change 49.7 meters, SD 68.8m, $P<.001$), despite relatively high levels of performance at baseline. Statistically significant improvements were also reported for the MacNew QOL questionnaire (mean change 0.28, SD 0.4, $P<.001$). We also

observed statistically significant improvements in the proportion of participants consuming 5 portions of fruit and vegetables each day: 22 (71%) to 29 (94%), $P=.01$. The proportion of those consuming at least 2 portions of oily fish each week rose from 14 (45%) to 21 (68%), $P=.01$. Anxiety and depression did not change significantly. There were no significant changes to the number of patients exercising for at least 30 minutes, 5 days per week: 27 (82%) pre and 26 (79%) post. No significant changes to smoking habits or BMI were observed in the study period.

Feedback from this cohort, indicated that 22 (54%) of the 41 recruited would not have attended a traditional out-patient CR program. Participants logged on to the program between 5 and 42 times, with an average of 10 times per participant.

Table 1. Baseline characteristics for the participants in the pilot study (n=41).

Characteristics	Value
Age in years, mean (SD)	60.5 (11.1)
Gender, m:f	37:4
Ethnicity, n (%)	
White/British	34 (82.9)
Indian	7 (17.1)
Initiating event, n (%)	
PCI	22 (53.7)
CABG	10 (24.4)
Other	2 (4.9)
Valve Surgery	4 (14.6)
Medical management	3 (7.3)
Weight in kg, mean (SD)	85.9 (12.1)
BMI in kg/m ² , mean (SD)	28.71 (5.9)

Table 2. Changes in exercise performance and health status (n=33).

	Pre, mean (SD)	Post, mean (SD)	Mean change (95% CI)	P value	Effect size
HADS anxiety	3.7 (2.9)	4.1 (3.2)	0.4 (-0.5 to 1.2)	.41	.1
HADS depression	2.5 (2.5)	2.6 (2.6)	0.1 (-0.5 to -0.7)	.84	.0
ISWT distance	580.6 (182.9)	630.3 (178.4)	49.7 (25.3-74.1)	<.001	.7
MacNew total	5.8 (0.8)	6.1 (0.7)	0.3 (0.1-0.4)	<.001	.7
Weight, kg	85.9 (12.8)	86.2 (12.1)	0.2 (-1.3 to 1.8)	.77	.1
BMI, kg/m ²	28.9 (6.3)	27.8 (3.2)	-1.2 (-3.3 to 1.0)	.29	.2

Table 3. Changes in health behaviors (n=33).

Behaviors	Pre	Post	P value
Exercise behavior, n (%)	27 (82)	26 (79)	1.0
Fruit and vegetables, n (%)	22 (71)	29 (94)	.01
Oily fish, n (%)	14 (45)	21 (68)	.01
Smoking, n (%)	1 (3)	0 (0)	1.0

Discussion

Principal Findings

This observational study demonstrates the potential benefits of using AYH as a form of CR to improve exercise capacity, quality of life, and dietary behavior in a low-risk population presenting for CR. The baseline exercise performance of this group was high, as expected in this cohort and their scope to increase their performance would be limited compared to those with lower baseline levels. The level of anxiety and depression were low at baseline, reflecting our selection criteria, again limiting the scope of improvement.

AYH may give patients a real choice of alternatives to supervised CR. Studies have shown that most patients with CHD who were still working preferred to follow a home-based rehab program to the conventional supervised CR program [5]. Alternative formats of CR, including home-based CR, such as the Heart Manual and the Angina Plan [38], have been shown to be an effective alternative to conventional CR [4]. However, these options are not widely available (5% and 2% of patients respectively in the United Kingdom [39]).

Patients require a CR program that is flexible to suit their lifestyle [40]. Using the Internet is a novel approach, providing flexibility for patients. It would potentially attract individuals who would otherwise be unable or unwilling to attend conventional CR, as traveling and the need to take time off work to attend CR classes is removed [25]. Important feedback from our patients indicated that over half (22/41, 54%) of the cohort would not have attended any other CR program. This demonstrates that there is a legitimate role for alternative formats of CR, such as AYH. Online CR has also shown to be a safe way to deliver CR [28].

Adherence to Web-based programs can be a potential issue as highlighted in previous studies [23]. Reasons for the non-compliance can vary from lack of time, to refusing to complete the program. However, several features have been identified that could help to improve adherence to a Web-based program. Making the program tailored to the user and interactive [41] as well as allowing users to set their own personal goals [42] have been shown to help improve adherence to Web-based interventions. The AYH program has incorporated all these features. The program is tailored to the individual, identifying their specific risk factors for CHD, providing interaction and self-monitoring. Self-monitoring is particularly beneficial as it has been shown to help increase an individual's awareness of their condition [43], thus allowing them to take better control of their condition.

In order for any Web-based program to be successful, it must suit the working practices of the health care professional (HCP) [44]. This was acknowledged when developing the AYH

program, ensuring that it can be easily incorporated into the CR service and does not increase the workload of the HCP. Each member of the CR team was provided with their own password to log onto an administration section of the program, allowing them to enroll patients on to the program, monitor their progress, and follow up with those who were not logging on or not progressing, via either email or phone call. The private messaging facility on the AYH program allowed for an easy and convenient way for the patient to access a CR specialist and vice versa. The ability to easily communicate with a large number of patients could have the added benefit of saving time and cost [45].

Following the pilot study, AYH has been incorporated into the CR service and made available to patients referred to the CR department. The criteria for inclusion were amended to allow for moderate risk patients to take part, that is, patients who completed at least level 5 (250 meters) of the ISWT and who had normal to moderate left ventricular dysfunction. We have collected data from 106 participants, using the outcomes employed in this study, which showed that after following AYH, there were statistically significant changes (mean change 47.16 m, $P<.001$) in exercise capacity, QOL ($P<.001$), and in a larger group, improvements in the number of patients exercising for at least 30 minutes, 5 days per week (increasing from 69.8% to 84.0%, $P=.001$). Statistically significant improvements were also noted for anxiety ($P=.01$) and depression ($P=.03$) and for those eating at least 2 portions of oily fish per week ($P=.02$). However, no changes were reported for consumption of fruit and vegetables or smoking habits. Feedback from this cohort indicated that over two-thirds of the participants (69/106, 65%) would not have attended a traditional out-patient CR program.

Limitations

The data reported are our initial findings from the AYH program; we now need to proceed to a clinical effectiveness trial incorporating a control group. Furthermore, the program was not offered to high-risk patients. Participants recorded low anxiety and depression scores and relatively high ISWT at baseline; nevertheless, statistically significant improvements were still observed.

In order to make AYH more accessible, a mobile version of AYH has also been developed. This allows users to access the program on their mobile device, such as smartphones and tablets (Figures 4 and 5). Using mobile phones to access CR has been shown to motivate patients to achieve their goals and to increase adherence [10]. A recent study [46] illustrated how sending emails to users helped increase adherence to the program. The latest version of AYH now creates automatic emails that are sent based on a user's level of interaction, for example, those not progressing or not logging on to the program for more than 7 days are flagged and emails are automatically sent to them.

Figure 4. Mobile version of the Activate Your Heart dashboard.

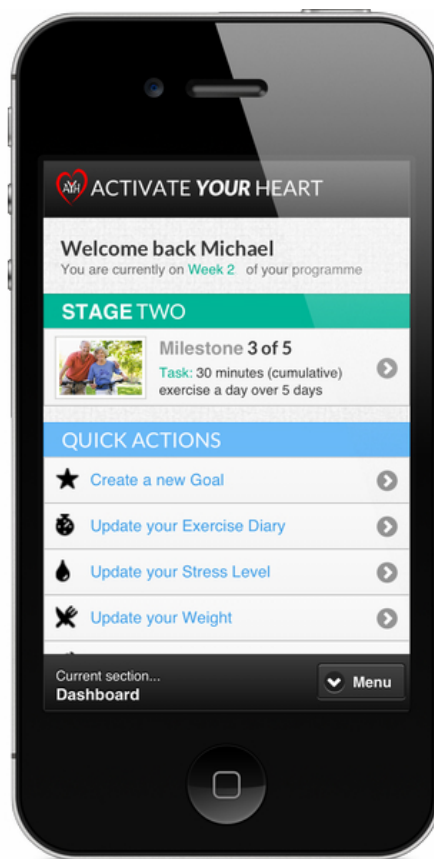
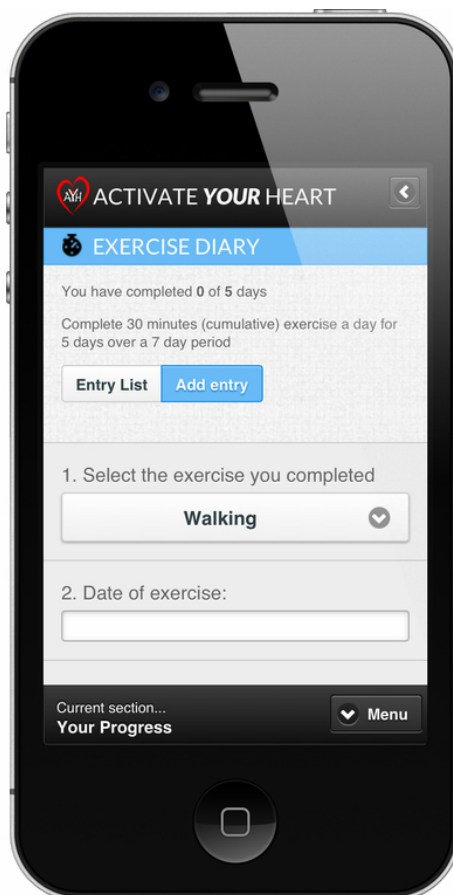


Figure 5. Adding exercise using a smartphone.



Conclusions

AYH has been designed to support individuals with CHD, promoting an alternative form of CR and provides a viable alternative option for patients, as well as providing a timely and patient-centered approach to CR. AYH enables patients to choose a program that best suits their lifestyle. The convenience

of using AYH means that many of the barriers associated with conventional CR programs are removed. To our knowledge, AYH is the first online comprehensive CR program to be introduced into a clinical service. This paper demonstrates that AYH can positively influence exercise capacity, QOL, and dietary behavior in a low-risk group and can be integrated into an existing CR service.

Acknowledgments

The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) in Leicestershire, Northamptonshire and Rutland and CLAHRC East Midlands and took place at the University Hospitals of Leicester NHS Trust. The authors would also like to acknowledge the Cardiac Rehabilitation team at the University Hospitals of Leicester NHS Trust for data collection and support. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Conflicts of Interest

None declared.

References

1. World Health Organization. World health organization fact sheet: Top 10 causes of death. URL: <http://www.who.int/mediacentre/factsheets/fs310/en/index.html> [accessed 2013-09-18] [WebCite Cache ID 6JisOGiP7]
2. Dalal HM, Zawada A, Jolly K, Moxham T, Taylor RS. Home based versus centre based cardiac rehabilitation: Cochrane systematic review and meta-analysis. *BMJ* 2010;340:b5631 [FREE Full text] [Medline: 20085991]
3. World Health Organization. Needs and action priorities in cardiac rehabilitation and secondary prevention in patients with CHD. 1993. URL: http://whqlibdoc.who.int/euro/-1993/EUR_ICP_CVD_125.pdf [accessed 2014-01-30] [WebCite Cache ID 6N0yUd7c9]
4. Jolly K, Lip GY, Taylor RS, Raftery J, Mant J, Lane D, et al. The Birmingham Rehabilitation Uptake Maximisation study (BRUM): a randomised controlled trial comparing home-based with centre-based cardiac rehabilitation. *Heart* 2009 Jan;95(1):36-42. [doi: 10.1136/hrt.2007.127209] [Medline: 18332063]
5. Grace SL, McDonald J, Fishman D, Caruso V. Patient preferences for home-based versus hospital-based cardiac rehabilitation. *J Cardiopulm Rehabil* 2005;25(1):24-29. [Medline: 15714108]
6. Thompson DR, Clark AM. Cardiac rehabilitation: into the future. *Heart* 2009 Dec;95(23):1897-1900. [doi: 10.1136/hrt.2009.173732] [Medline: 19815940]
7. Clark AM, King-Shier KM, Spaling MA, Duncan AS, Stone JA, Jaglal SB, et al. Factors influencing participation in cardiac rehabilitation programmes after referral and initial attendance: qualitative systematic review and meta-synthesis. *Clin Rehabil* 2013 Oct;27(10):948-959. [doi: 10.1177/0269215513481046] [Medline: 23798748]
8. De Vos C, Li X, Van Vlaenderen I, Saka O, Dendale P, Eyssen M, et al. Participating or not in a cardiac rehabilitation programme: factors influencing a patient's decision. *Eur J Prev Cardiol* 2013 Apr;20(2):341-348. [doi: 10.1177/2047487312437057] [Medline: 22345682]
9. Varnfield M, Karunanithi MK, Särelä A, Garcia E, Fairfull A, Oldenburg BF, et al. Uptake of a technology-assisted home-care cardiac rehabilitation program. *Med J Aust* 2011 Feb 21;194(4):S15-S19. [Medline: 21401482]
10. Sharp J, Freeman C. Patterns and predictors of uptake and adherence to cardiac rehabilitation. *J Cardiopulm Rehabil Prev* 2009;29(4):241-247. [doi: 10.1097/HCR.0b013e3181adcf0f] [Medline: 19628988]
11. Blair J, Corrigan H, Angus NJ, Thompson DR, Leslie S. Home versus hospital-based cardiac rehabilitation: a systematic review. *Rural Remote Health* 2011;11(2):1532 [FREE Full text] [Medline: 21488706]
12. Jolly K, Taylor R, Lip G, Greenfield S, Raftery J, Mant J, et al. The Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Home-based compared with hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence. *Health Technol Assess* 2007 Sep;11(35):1-118 [FREE Full text] [Medline: 17767899]
13. Oerkild B, Frederiksen M, Hansen JF, Prescott E. Home-based cardiac rehabilitation is an attractive alternative to no cardiac rehabilitation for elderly patients with coronary heart disease: results from a randomised clinical trial. *BMJ Open* 2012;2(6):- [FREE Full text] [doi: 10.1136/bmjopen-2012-001820] [Medline: 23253876]
14. Dutton WH, Blank G. Oxford Internet Survey 2013 Report. URL: http://oxis.oii.ox.ac.uk/sites/oxis.oii.ox.ac.uk/files/content/files/publications/OxIS_2013.pdf [accessed 2014-01-31] [WebCite Cache ID 6N2BgusQI]
15. Office of National Statistics. Internet Access Quarterly Update. URL: http://www.ons.gov.uk/ons/dcp171778_336739.pdf [accessed 2014-10-10] [WebCite Cache ID 6TDmL6hs]

16. 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](https://doi.org/10.1016/j.pec.2007.01.009)] [Medline: [17329062](https://pubmed.ncbi.nlm.nih.gov/17329062/)]
17. 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](https://doi.org/10.2196/jmir.1035)] [Medline: [19275980](https://pubmed.ncbi.nlm.nih.gov/19275980/)]
18. Kraal JJ, Peek N, van den Akker-Van Marle ME, Kemps HM. Effects and costs of home-based training with telemonitoring guidance in low to moderate risk patients entering cardiac rehabilitation: The FIT@Home study. *BMC Cardiovasc Disord* 2013;13:82 [FREE Full text] [doi: [10.1186/1471-2261-13-82](https://doi.org/10.1186/1471-2261-13-82)] [Medline: [24103384](https://pubmed.ncbi.nlm.nih.gov/24103384/)]
19. Kuhl EA, Sears SF, Conti JB. Internet-based behavioral change and psychosocial care for patients with cardiovascular disease: a review of cardiac disease-specific applications. *Heart Lung* 2006;35(6):374-382. [doi: [10.1016/j.hrtlng.2006.02.004](https://doi.org/10.1016/j.hrtlng.2006.02.004)] [Medline: [17137938](https://pubmed.ncbi.nlm.nih.gov/17137938/)]
20. Schulz DN, Kremers SP, van Osch LA, Schneider F, van Adrichem MJ, de Vries H. Testing a Dutch web-based tailored lifestyle programme among adults: a study protocol. *BMC Public Health* 2011;11:108 [FREE Full text] [doi: [10.1186/1471-2458-11-108](https://doi.org/10.1186/1471-2458-11-108)] [Medline: [21324181](https://pubmed.ncbi.nlm.nih.gov/21324181/)]
21. 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](https://doi.org/10.2196/jmir.6.4.e40)] [Medline: [15631964](https://pubmed.ncbi.nlm.nih.gov/15631964/)]
22. Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. *Med J Aust* 2010 Jun 7;192(11 Suppl):S4-11. [Medline: [20528707](https://pubmed.ncbi.nlm.nih.gov/20528707/)]
23. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. *Int J Behav Nutr Phys Act* 2009;6:30 [FREE Full text] [doi: [10.1186/1479-5868-6-30](https://doi.org/10.1186/1479-5868-6-30)] [Medline: [19490649](https://pubmed.ncbi.nlm.nih.gov/19490649/)]
24. Song M, Choe MA, Kim KS, Yi MS, Lee I, Kim J, et al. An evaluation of Web-based education as an alternative to group lectures for diabetes self-management. *Nurs Health Sci* 2009 Sep;11(3):277-284. [doi: [10.1111/j.1442-2018.2009.00458.x](https://doi.org/10.1111/j.1442-2018.2009.00458.x)] [Medline: [19689636](https://pubmed.ncbi.nlm.nih.gov/19689636/)]
25. Vernooij JW, Kaasjager HA, van der Graaf Y, Wierdsma J, Grandjean HM, Hovens MM, SMARTStudy Group. Internet based vascular risk factor management for patients with clinically manifest vascular disease: randomised controlled trial. *BMJ* 2012;344:e3750 [FREE Full text] [Medline: [22692651](https://pubmed.ncbi.nlm.nih.gov/22692651/)]
26. Reid RD, Morrin LI, Beaton LJ, Papadakis S, Kocourek J, McDonnell L, et al. Randomized trial of an internet-based computer-tailored expert system for physical activity in patients with heart disease. *Eur J Prev Cardiol* 2012 Dec;19(6):1357-1364. [doi: [10.1177/1741826711422988](https://doi.org/10.1177/1741826711422988)] [Medline: [21903744](https://pubmed.ncbi.nlm.nih.gov/21903744/)]
27. Southard BH, Southard DR, Nuckolls J. Clinical trial of an Internet-based case management system for secondary prevention of heart disease. *J Cardiopulm Rehabil* 2003;23(5):341-348. [Medline: [14512778](https://pubmed.ncbi.nlm.nih.gov/14512778/)]
28. Zutz A, Ignaszewski A, Bates J, Lear SA. Utilization of the internet to deliver cardiac rehabilitation at a distance: a pilot study. *Telemed J E Health* 2007 Jun;13(3):323-330. [doi: [10.1089/tmj.2006.0051](https://doi.org/10.1089/tmj.2006.0051)] [Medline: [17603835](https://pubmed.ncbi.nlm.nih.gov/17603835/)]
29. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992 Dec;47(12):1019-1024 [FREE Full text] [Medline: [1494764](https://pubmed.ncbi.nlm.nih.gov/1494764/)]
30. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](https://pubmed.ncbi.nlm.nih.gov/6880820/)]
31. Hillers TK, Guyatt GH, Oldridge N, Crowe J, Willan A, Griffith L, et al. Quality of life after myocardial infarction. *J Clin Epidemiol* 1994 Nov;47(11):1287-1296. [Medline: [7722565](https://pubmed.ncbi.nlm.nih.gov/7722565/)]
32. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res* 2002 Feb;52(2):69-77. [Medline: [11832252](https://pubmed.ncbi.nlm.nih.gov/11832252/)]
33. Höfer S, Lim L, Guyatt G, Oldridge N. The MacNew Heart Disease health-related quality of life instrument: a summary. *Health Qual Life Outcomes* 2004;2:3 [FREE Full text] [doi: [10.1186/1477-7525-2-3](https://doi.org/10.1186/1477-7525-2-3)] [Medline: [14713315](https://pubmed.ncbi.nlm.nih.gov/14713315/)]
34. Jolly K, Taylor RS, Lip GY, Singh S, BRUM Steering Committee. Reproducibility and safety of the incremental shuttle walking test for cardiac rehabilitation. *Int J Cardiol* 2008 Mar 28;125(1):144-145. [doi: [10.1016/j.ijcard.2007.01.037](https://doi.org/10.1016/j.ijcard.2007.01.037)] [Medline: [17408777](https://pubmed.ncbi.nlm.nih.gov/17408777/)]
35. Arnold HJ, Sewell L, Singh S. A comparison of once-versus twice-weekly supervised phase III cardiac rehabilitation. *Br J Cardiol* 2007;14:45-48.
36. Harrison SL, Greening NJ, Houchen-Wolloff L, Bankart J, Morgan MD, Steiner MC, et al. Age-specific normal values for the incremental shuttle walk test in a healthy British population. *J Cardiopulm Rehabil Prev* 2013;33(5):309-313. [doi: [10.1097/HCR.0b013e3182a0297e](https://doi.org/10.1097/HCR.0b013e3182a0297e)] [Medline: [23959208](https://pubmed.ncbi.nlm.nih.gov/23959208/)]
37. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
38. Lewin RJ, Furze G, Robinson J, Griffith K, Wiseman S, Pye M, et al. A randomised controlled trial of a self-management plan for patients with newly diagnosed angina. *Br J Gen Pract* 2002 Mar;52(476):194-6, 199 [FREE Full text] [Medline: [12030661](https://pubmed.ncbi.nlm.nih.gov/12030661/)]
39. NACR. The National Audit of Cardiac Rehabilitation: Annual Statistical Report. 2013. URL: <http://www.cardiacrehabilitation.org.uk/nacr/docs/2013.pdf> [accessed 2014-01-28] [WebCite Cache ID 6MxeSGxup]

40. Pollard M, Sutherland C. Cardiac rehabilitation: Are we putting our hearts into it? *Br J Cardiol* 2009;16(5):247-249.
41. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. *Med 2.0* 2012 Aug 14;1(2):e3. [doi: [10.2196/med20.1741](https://doi.org/10.2196/med20.1741)]
42. Vandelanotte C, Dwyer T, Van Itallie A, Hanley C, Mummery WK. The development of an internet-based outpatient cardiac rehabilitation intervention: a Delphi study. *BMC Cardiovasc Disord* 2010;10:27 [FREE Full text] [doi: [10.1186/1471-2261-10-27](https://doi.org/10.1186/1471-2261-10-27)] [Medline: [20534167](https://pubmed.ncbi.nlm.nih.gov/20534167/)]
43. Goessens BM, Visseren FL, de Nooijer J, van den Borne HW, Algra A, Wierdsma J, SMART Study Group. A pilot-study to identify the feasibility of an Internet-based coaching programme for changing the vascular risk profile of high-risk patients. *Patient Educ Couns* 2008 Oct;73(1):67-72. [doi: [10.1016/j.pec.2008.06.004](https://doi.org/10.1016/j.pec.2008.06.004)] [Medline: [18639410](https://pubmed.ncbi.nlm.nih.gov/18639410/)]
44. 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](https://doi.org/10.2196/resprot.1804)] [Medline: [23612470](https://pubmed.ncbi.nlm.nih.gov/23612470/)]
45. Weinstein PK. A review of weight loss programs delivered via the Internet. *J Cardiovasc Nurs* 2006;21(4):251-8; quiz 259. [Medline: [16823276](https://pubmed.ncbi.nlm.nih.gov/16823276/)]
46. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. *J Med Internet Res* 2013;15(1):e23 [FREE Full text] [doi: [10.2196/jmir.2151](https://doi.org/10.2196/jmir.2151)] [Medline: [23363466](https://pubmed.ncbi.nlm.nih.gov/23363466/)]

Abbreviations

AYH: Activate Your Heart
BMI: body mass index
CABG: coronary artery bypass graft
CHD: coronary heart disease
CR: cardiac rehabilitation
HADS: Hospital Anxiety Depression Score
HCP: health care professional
ISWT: Incremental Shuttle Walk Test
PCI: percutaneous coronary intervention
QOL: Quality of Life

Edited by G Eysenbach; submitted 18.10.13; peer-reviewed by P Doherty, S Fraser; comments to author 15.11.13; revised version received 31.01.14; accepted 31.05.14; published 29.10.14.

Please cite as:

Brough C, Boyce S, Houchen-Wolloff L, Sewell L, Singh S
Evaluating the Interactive Web-Based Program, Activate Your Heart, for Cardiac Rehabilitation Patients: A Pilot Study
J Med Internet Res 2014;16(10):e242
URL: <http://www.jmir.org/2014/10/e242/>
doi: [10.2196/jmir.3027](https://doi.org/10.2196/jmir.3027)
PMID: [25359204](https://pubmed.ncbi.nlm.nih.gov/25359204/)

©Christopher Brough, Sally Boyce, Linzy Houchen-Wolloff, Louise Sewell, Sally Singh. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Comparison Between Phone-Based Psychotherapy With and Without Text Messaging Support In Between Sessions for Crisis Patients

Gareth Furber^{1*}, PhD (Clin Psycho); Gabrielle Margaret Jones^{2*}, BHLthSc; David Healey^{3*}, BSW, MMHSc; Niranjan Bidargaddi^{4,5*}, BE (CS), PhD (bioinformatics)

¹Health Economics and Social Policy Group, School of Population Health, University of South Australia, Adelaide, Australia

²Mental Health Observatory Research Unit, School of Medicine, The University of Adelaide, Adelaide, Australia

³IAPT@Flinders unit, Southern Mental Health, Flinders Medical Centre, Adelaide, Australia

⁴Mental Health Informatics Research Unit, Country Health SA LHN Inc, Adelaide, Australia

⁵e-Health Research, School of Medicine, Flinders University, Adelaide, Australia

* all authors contributed equally

Corresponding Author:

Gareth Furber, PhD (Clin Psycho)

Health Economics and Social Policy Group

School of Population Health

University of South Australia

South Australian Health & Medical Research Institute (SAHMRI)

North Terrace

Adelaide, 5000

Australia

Phone: 61 80429344065

Fax: 61 83022337

Email: gareth.furber@unisa.edu.au

Abstract

Background: Few studies have tested whether individually tailored text messaging interventions have an effect on clinical outcomes when used to supplement traditional psychotherapy. This is despite the potential to improve outcomes through symptom monitoring, prompts for between-session activities, and psychoeducation.

Objective: The intent of the study was to explore the use of individually tailored between-session text messaging, or short message service (SMS), as an adjunct to telephone-based psychotherapy for consumers who present to the Emergency Department (ED) in situational and/or emotional crises.

Methods: Over a 4-month period, two therapists offered 68 prospective consumers of a telephone-based psychotherapy service individually tailored between-session text messaging alongside their telephone-based psychotherapy. Attendance and clinical outcomes (depression, anxiety, functional impairment) of those receiving messages were compared against a historical control group (n=157) who received telephone psychotherapy only.

Results: A total of 66% (45/68) of the consumers offered SMS accepted the intervention. A total of 432 messages were sent over the course of the trial, the majority involving some kind of psychoeducation or reminders to engage in therapy goals. There were no significant differences in clinical outcomes between consumers who received the SMS and those in the control group. There was a trend for participants in the intervention group to attend fewer sessions than those in the control group (mean 3.7, SD 1.9 vs mean 4.4, SD 2.3).

Conclusions: Both groups showed significant improvement over time. Individually tailored SMS were not found to improve clinical outcomes in consumers receiving telephone-based psychotherapy, but the study was underpowered, given the effect sizes noted and the significance level chosen. Given the ease of implementation and positive feedback from therapists and clients, individually tailored text messages should be explored further in future trials with a focus on enhancing the clinical impact of the tailored text messages, and utilizing designs with additional power to test for between-group effects.

KEYWORDS

telemedicine; psychotherapy; mental health services; mobile health; mHealth; short message service; eHealth

Introduction

Multiple authors have discussed the value of mobile phones in the treatment of mental disorders [1-8]. The characteristics of mobile phones—popularity, widespread use, low cost, portable, always on, advanced multimedia capabilities, can run complex apps—mean there are many avenues for mental health professionals to harness their potential in treatment. Despite the significant potential, only a small number of studies have tested whether supplementing mental health care with mobile phone-based interventions leads to better consumer outcomes.

Three studies have shown that symptom monitoring alongside usual care can lead to better outcomes. Spaniel and colleagues [9] used Short Message Service (SMS), or text messaging, for symptom monitoring to identify prodromal symptoms in patients with schizophrenia and alert psychiatrists to initiate medication changes. The addition of SMS monitoring to their usual case management was shown to reduce the number of hospitalizations and hospitalization length. Kramer et al [10] used a system of mobile phone symptom monitoring and clinical feedback in patients receiving medication for depression. Patients who were monitored and provided feedback showed a larger decrease in depression symptoms over 6-month follow-up, compared to both a control (no monitoring) and pseudo-control (monitoring but no feedback). Reid, Kauer, and colleagues [11,12] found that a daily mobile phone mental health assessment and feedback program delivered through primary care to young people with mental health problems, led to improvements in emotional self-awareness and depression.

Two studies have shown that the use of text message reminders alongside usual care can improve treatment processes. Pijnenborg et al [13] used daily text message reminders as a form of “cognitive aid” to assist patients with schizophrenia in achieving treatment goals (eg, attend appointments, take their medication). When receiving reminders, patients showed a significant increase in percentage of goals achieved. Similarly, Montes and colleagues [14] showed that daily text message medication reminders for patients with schizophrenia increased medication adherence compared to those who did not receive reminders.

Finally, three studies have shown that delivering mobile phone-based aftercare can improve post-treatment outcomes. Bauer and colleagues [15] found that a text messaging-based symptom reporting and tailored feedback system delivered post-treatment to women with eating disorders led to increased remission rates at 8 months post-treatment. Agyapong et al [16] found twice-daily supportive/educational text messages relating to mood and alcohol abstinence led to lower depression scores at 3 months post-discharge in patients who had recently completed an in-patient dual diagnosis (depression and alcohol use) treatment program. Marasinghe and colleagues [17] demonstrated that a brief mobile phone-based treatment

involving training and advice on problem solving, social support, and drugs and alcohol, delivered to patients following hospitalization for a suicide attempt, led to lower rates of suicidal ideation and depression at 6 and 12 months post-hospitalization.

In all of these examples, mobile phones were used to supplement standard care (ie, case management, primary care, medication). The evidence that mobile phone apps can improve outcomes of psychotherapy is less convincing. Case studies and small uncontrolled trials have demonstrated that mobile devices (including mobile phones) can feasibly be used as part of psychotherapy such as cognitive behavior therapy (CBT). For example, personal digital assistants (PDA), precursors to modern smartphones, have been used to assist in the cognitive behavioral treatment of obsessive compulsive disorder [18], schizophrenia [19], social phobia [20], bulimia nervosa [21], and generalized anxiety disorder [22]. More recently, Aguilera et al [23] used text messaging to send participants attending a group CBT program, daily messages regarding mood monitoring and content of group sessions. Participants reported the messages made them feel closer to the group and more likely to attend group sessions. Rizvi and colleagues have also piloted the use of a mobile phone app to complement dialectical behavior treatment for borderline personality disorder [24]. However, none of these studies used control groups to determine if the addition of the mobile phone improved consumer outcomes.

The goal of the current study was to develop an individualized text messaging support intervention as an adjunct to an existing telephone-based psychotherapy service and evaluate it using a historical control group. The setting for this study was a small team of psychotherapists who provide telephone-based psychotherapy to consumers who recently presented to the Emergency Department (ED) in situational and/or emotional crises. To our knowledge, the value of supplementing psychotherapy in this population with tailored text messaging interventions has not been explored.

We developed a text messaging intervention where therapists and consumers collaborate to compose goal reminder and psychoeducation messages, individually tailored to the consumer’s goals, and deliver these messages between sessions. This was in response to emerging evidence that individuals are sensitive to variations in the linguistic content of text messages designed to help them achieve a personal goal [25]. Tailoring text message content to an individual can be more effective compared to automated pre-determined messages [26]. The choice to develop a text messaging intervention (rather than a mobile phone app) was made because text messaging is widely available on mobile phones, has an extremely low learning curve, and requires minimally expensive and complex infrastructure. We also noted that the majority of work in the area, as described above, has used text messaging, and that reviews of text messaging interventions in the wider health

literature (eg, [27-29]) have suggested significant value in their use.

Methods

Setting

The current project was conducted within the “IAPT@Flinders” service. IAPT@Flinders is an Australian outgrowth of the United Kingdom’s Improving Access to Psychological Therapies (IAPT) model, which provides rapid access to evidence-based, low intensity psychological interventions, based on CBT principles for people suffering mild to moderate depression and anxiety [30]. IAPT@Flinders targets consumers who present to the ED in situational and/or emotional crises who are not subsequently admitted to psychiatric services. The fundamental goal of the IAPT@ Flinders service is to provide an alternative evidence-based rapid access treatment pathway for these consumers, and thus reduce the burden on ED of mental health crisis readmissions. Inclusion criteria for the service include: medical clearance to leave the ED, aged 18 years and over, not currently engaged with a mental health service or currently receiving psychological therapy, and experiencing strong psychological distress including anxiety and/or depressive symptoms.

The service provided to IAPT@Flinders consumers includes low intensity telephone-based CBT as well as referral and linkages to other key community services (eg, relationship, drug and alcohol, employment, housing, or financial services). The low-intensity CBT treatment component consists of providing consumers with self-help resources relevant to their presenting symptoms and supporting them in working through these materials. The self-help resources (eg, [31]) typically include psychoeducation about depression and/or anxiety and a range of cognitive and behavioral strategies such as behavioral activation, cognitive restructuring, exposure therapy, sleep and panic management, and motivational problem solving. The telephone sessions involve explaining key CBT principles and supporting consumers in setting goals based on the techniques described in the self-help materials. A unique aspect of the IAPT program that facilitates evaluation is the use of session-by-session measures of outcome. This means any consumer attending at least two appointments will have two data points.

Study Design and Participants

We hypothesized that consumers who received individually tailored text messaging in addition to their psychotherapy would show improved clinical outcomes (levels of depression, anxiety, and functional impairment) and increased attendance compared to those receiving psychotherapy alone.

Two IAPT@Flinders therapists from a pool of four volunteered to take part in the trial after it was described to them by the authors (GF, GJ, NB). Over a 4-month period in 2013, these two therapists offered all consumers presenting to the service the opportunity to receive individually tailored text messages alongside their usual care. For a consumer to be offered the text messaging, they needed to (1) meet criteria for treatment with IAPT@Flinders, (2) attend the initial assessment (telephone or

face-to-face), (3) consent to receive text messages, and (4) own and be able to use a mobile phone that could receive text messages. Consumers declining the text messaging intervention were not evaluated further. The sample was considered a self-selected sample.

The impact of individually tailored text messaging was assessed by comparing attendance and clinical outcomes of those receiving the text messages (intervention group) to a historical control group. The choice to use a historical control group versus control group by randomization was pragmatic, reflecting multiple barriers to implementing randomization in the service, including resources, personnel, risk of consumer disengagement, and contamination.

Intervention Group

The intervention group received individually tailored text messages alongside their usual care. The text messaging intervention was developed collaboratively by IAPT@Flinders clinicians and the IAPT@Flinders research group (GF, GJ, NB), and was designed to complement the goal setting already being conducted with consumers. The intervention was designed to be simple to implement, low cost, and involve minimal extra clinician time. Two basic message types were conceptualized. The first (reminders) were specific prompts to engage in activities/homework discussed in session. For example, consumers were reminded to engage in physical activity if they had set that as a goal for the week. The second (psychoeducation) were reiterations of key concepts discussed during therapy. For example, consumers were reminded of recommended methods for managing feelings of anxiety such as slow breathing or reassurance that feelings of anxiety will pass naturally.

Participants were introduced to the intervention as follows. While in treatment and at the end of each session, the therapist would ask the consumer if it would be useful to send some text message reminders and psychoeducation during the upcoming week/fortnight in relation to what was discussed during the session. If the consumer consented, the therapist and consumer would collaboratively plan the content and schedule of those messages. As such, all messages were uniquely tailored to the consumer’s goals for that week/fortnight and the consumer was free to choose whether to receive text messages between each session.

Text messages were sent one-way and their delivery was coordinated by a Web-based program supplied by goACT Pty Ltd that enabled therapists to schedule messages for individual consumers, save message “templates” for use across different consumers, and to share between therapists. The content and schedule of messages was decided on a session-by-session basis and programmed into the system by the therapist at the end of the session. When reviewing consumers’ progress toward goals set in the previous week/fortnight, therapists inquired about the usefulness of the text messages in the previous week/fortnight.

Two safety protocols were developed for the study. The first required that all messages include the name of the therapist and the service: “Hi (first name of consumer), (first name of therapist) from IAPT here”. This was done to ensure consumers

knew who the messages were coming from. The second protocol instructed therapists to cease scheduled messages (ie, delete them from the program) should it become apparent that the consumer had been re-hospitalized for self-harm or suicide attempt. This was to reduce the likelihood of the consumer being sent inappropriate messages during a crisis episode.

Historical Control Group

Attendance and clinical outcomes of consumers receiving the text messaging intervention were compared to a historical control group. This group was made up of consumers who had received a service from IAPT@Flinders (from the same two therapists) in the 6 months prior to the introduction of the text messaging intervention. Data for the control group was extracted anonymously from the electronic clinical record. Similar to the criteria for the intervention group, only those who (1) met criteria for treatment with IAPT@Flinders, and (2) attended the initial assessment (telephone or face-to-face) and at least one other session were included in the historical control group. The control group consisted of 157 consumers.

Measures

At intake to the service, all consumers provided their age, gender, employment status (employed, unemployed), education (high school complete: yes/no), relationship status (in a relationship: yes/no), prescribed psychotropic medication (yes/no), and whether they had a chronic health condition (yes/no).

Attendance was measured by recording treatment length (in days), whether treatment was completed, the number of telephone and face-to-face sessions that a consumer attended, as well as the number of recorded “did not attend” (DNA) sessions.

Clinical outcomes were measured using the 9-item Patient Health Questionnaire (PHQ-9), the 7-item Generalized Anxiety Disorder questionnaire (GAD-7), and the Work and Social Adjustment Scale (WSAS). For all consumers, the “pre” and “post” clinical outcomes corresponded to the first and last recorded instances of the three measures during their episode of care. On average, the time between the first and last recorded measures was 6 weeks.

The PHQ-9 is a 9-item self-reported questionnaire designed to evaluate the presence of depressive symptoms during the prior 2 weeks [32]. Each of the 9 items, relating to each of the DSM-IV diagnostic criteria for depression, can be scored from 0 (not at all) to 3 (nearly every day). In IAPT@Flinders, the PHQ-9 is used as a severity measure, with scores ranging from 0 (absence of depressive symptoms) to 27 (severe depressive symptoms).

The GAD-7 is a 7-item self-reported questionnaire designed to assess the presence of the symptoms of Generalized Anxiety Disorder (GAD), as listed in the DSM-IV [33]. The total GAD-7 score is calculated by summing the scores for all 7 items, which can be scored from 0 (Not at all) to 3 (nearly every day). In IAPT@Flinders, the GAD-7 is used as a severity measure, with scores ranging from 0 (minimal) to 21 (serious).

The WSAS is a 5-item self-reported questionnaire that assesses the degree of functional impairment attributable to an identified problem [34]. The WSAS score is calculated by summing the scores for all 5 items, which range from 0 (not at all) to 8 (severely). In IAPT@Flinders, the WSAS is used as an indicator of impairment with scores ranging from 0 (no impairment) to 40 (severe impairment).

Analysis

T tests (continuous variables) and chi-square tests (categorical variables) were used to test for differences between the intervention and control groups on measures taken at intake and measures of attendance. Repeated measures analysis of variance (ANOVA) was used to compare the intervention and control groups on change in clinical outcomes (PHQ, GAD, and WSAS) over time. Analyses were conducted in SPSS Version 19.

Results

Of the 111 consumers allocated to the two participating therapists during the study period, 68 were offered the text messaging intervention. Reasons for exclusion are outlined in [Figure 1](#). Therapist #2 was successful in recruiting 86% (31/36) of 36 consumers offered the intervention, in contrast with a consent rate of 44% (14/32) achieved by Therapist #1. In total, 45 consumers consented to and received the text messaging intervention alongside their usual IAPT@Flinders care. The historical control group was made up of 157 of the 232 clients recorded in the system at the time of analysis ([Figure 2](#)).

Intake status, attendance, and clinical outcomes for the intervention and control groups are summarized in [Table 1](#). On a number of indicators, the intervention group appeared to be functioning better at intake: more were employed, more were in a relationship, fewer were prescribed psychotropic medication, and fewer reported a chronic health condition. Only one of these differences was statistically significant—consumers in the intervention group were less likely to be prescribed a psychotropic medication. Intervention and control groups did not differ significantly on their intake PHQ, GAD, or WSAS scores, suggesting similar levels of psychological distress.

Consumers in the intervention group were sent a total of 432 messages during the trial, consisting of 150 unique messages. Of these 150, 39 messages were sent two or more times, with the remainder being sent once only. The bulk of message content related to psychoeducation (eg, “Anxiety is normal and it will pass with time. Remember to slow your breathing when you start to feel anxious”), and reminders to engage in discussed homework tasks (eg, “Remember to schedule and complete behavioral activation tasks”).

Therapist #2 was responsible for the bulk of the messages sent (87.7%, 378/431). On average, consumers were sent 9.6 messages each, although there was wide variation. Over 90% (94.2%, 406/431) of messages were sent within 4 weeks of the consumers commencing therapy, with over 60% (61.9%, 267/431) sent in the first two weeks. Consistent with this, therapists reported that consumers tended to find the messages useful early in therapy, and less useful over time.

Regarding attendance, there were no statistically significant differences between intervention and control groups on treatment completion, treatment length, number of sessions attended, or DNA rate. A trend ($P=.063$) was observed for intervention participants to attend fewer sessions ($-.7$) than control participants. To explore this further, we charted the proportion of intervention and control groups by number of sessions received (Figure 3). This revealed that a greater proportion of intervention participants received 1-4 sessions whereas a greater proportion of control participants received 5-12 sessions.

Regarding clinical outcomes, there was a non-significant main effect of group on PHQ scores, $F_{1,199}=0.090$, $P=.764$, partial $\eta^2=0.000$, indicating intervention and control groups did not differ in their overall level of depressive symptomatology. There was a significant main effect of time on PHQ scores, $F_{1,199}=336$, $P=.000$, partial $\eta^2=0.629$, indicating both groups showed improvements in depressive symptomatology over time. There was a non-significant interaction between group and time on PHQ scores, $F_{1,199}=1.065$, $P=.303$, partial $\eta^2=0.005$, indicating no differences between groups in how depressive symptomatology changed over time.

There was a non-significant main effect of group on GAD scores, $F_{1,199}=0.508$, $P=.477$, partial $\eta^2=0.003$, indicating intervention and control groups did not differ in their level of anxiety symptomatology. There was a significant main effect of time on GAD scores, $F_{1,199}=283$, $P=.000$, partial $\eta^2=0.587$, indicating both groups showed improvements in anxiety symptomatology over time. There was a non-significant interaction between group and time on GAD scores,

$F_{1,199}=2.346$, $P=.127$, partial $\eta^2=0.012$, indicating no differences between groups in how anxiety symptomatology changed over time.

There was a non-significant main effect of group on WSAS scores, $F_{1,199}=0.135$, $P=.713$, partial $\eta^2=0.001$, indicating intervention and control groups did not differ in their level of functional impairment. There was a significant main effect of time on WSAS scores, $F_{1,199}=199$, $P=.000$, partial $\eta^2=0.500$, indicating both groups showed improvements in functional impairment over time. There was a non-significant interaction between group and time on WSAS scores, $F_{1,199}=199$, $P=.892$, partial $\eta^2=0.004$, indicating no differences between groups in how functional impairment scores changed over time.

Additional ANOVA were conducted in which only those participants seen by Therapist #2 were included and where prescribed medication was included as a covariate. The pattern of results for PHQ, GAD, and WSAS did not change under these alternative models.

To determine if the study was adequately powered to detect differences between groups on attendance and clinical outcomes, we used G-Power [35] to conduct two post-hoc power analyses for the outcomes of number of sessions attended and change in PHQ scores over time (Table 2). These analyses revealed the probability of rejecting the null hypothesis, that the intervention and control groups were the same, was low for both number of sessions attended (.499), and change in PHQ scores over time (.378). Analyses of differences between groups should therefore be treated with caution.

Table 1. Intake status, attendance, and clinical outcomes for the intervention and control groups.

	Intervention group (n=45)	Control group (n=157)	Difference (95% CI)	Comparisons (t test, Pearson chi-square)	P value
Intake status					
Age at assessment, mean (SD)	34.1 (14.5)	37 (15.8)	-3.51 (-8.7 to 1.67)	$t_{197}=-1.34$	$P=.183$
Gender, n (%) male	15 (33.3)	68 (43.3)	-10%	$\chi^2_1=1.44$	$P=.230$
Employment, n (%) employed	22 (48.9)	68 (43.3)	5.6%	$\chi^2_1=0.44$	$P=.507$
Relationship, n (%) yes	18 (40.0)	44 (28.0)	12%	$\chi^2_1=2.12$	$P=.145$
High school complete, n (%) yes	20 (44.4)	64 (40.8)	3.6%	$\chi^2_1=2.28$	$P=.131$
Prescribed psychotropic medication, n (%) yes	24 (53.3)	113 (72)	-18.7%	$\chi^2_1=6.02$	$P=.014$
Chronic health condition, n (%) yes	13 (28.9)	55 (35.0)	-6.1%	$\chi^2_1=0.44$	$P=.509$
Attendance					
Treatment completers, n (%)	28 (62.2)	111 (70.7)	-8.5%	$\chi^2_1=1.17$	$P=.279$
Treatment length in days, mean (SD)	43.3 (17.1)	46.3 (26.7)	-3.05 (-11.74 to 5.64)	$t_{194}=-0.89$	$P=.376$
Number of sessions, mean (SD)	3.73 (1.9)	4.43 (2.3)	-0.700 (-1.44 to 0.04)	$t_{200}=-1.87$	$P=.063$
Number of DNA ^a , mean (SD)	1.42 (1.12)	1.76 (1.51)	-0.342 (-0.82 to 0.14)	$t_{200}=-1.66$	$P=.100$
Clinical outcomes					
PHQ ^b Pre, mean (SD)	19.4 (6.4)	18.9 (5.6)	0.480 (-1.46 to 2.42)	$t_{200}=0.49$	$P=.626$
GAD ^c Pre, mean (SD)	16.1 (5.2)	15.8 (4.4)	0.350 (-1.18 to 1.88)	$t_{200}=0.45$	$P=.653$
WSAS ^d Pre, mean (SD)	27.7 (9.8)	27.0 (10.0)	0.730 (-0.26 to 4.06)	$t_{200}=0.43$	$P=.666$
PHQ Post, mean (SD)	6.64 (6.23)	7.57 (6.85)			
GAD Post, mean (SD)	5.41 (5.51)	6.81 (6.17)			
WSAS Post, mean (SD)	10.89 (10.70)	12.49 (12.44)			

^aDNA: did not attend^bPHQ: Patient Health Questionnaire^cGAD: Generalized Anxiety Disorder questionnaire^dWSAS: Work and Social Adjustment Scale

Table 2. Post hoc power analyses for number of sessions attended and PHQ^ascores.

	Number of sessions attended	PHQ scores
Test	<i>t</i> tests / Means: difference between two independent means (two groups)	<i>F</i> tests / ANOVA: Repeated measures, within-between interaction
Analysis	Post hoc: Compute achieved power	Post hoc: Compute achieved power
Inputs	Tail(s)=Two Effect size <i>d</i> =0.3326908 α err prob=0.05 Sample size group 1=157 Sample size group 2=45	Effect size <i>f</i> =0.0708881 α err prob=0.05 Total sample size=203 Number of groups=2 Number of measurements=2 Correlation among repeated measures=0.257 Nonsphericity correction ε=1
Output	Noncentrality parameter δ=1.9675292 Critical <i>t</i> =1.9718962 df=200 Power (1-β err prob)=0.4992920	Noncentrality parameter λ=2.7458948 Critical <i>F</i> =3.8881392 Numerator df=1.0000000 Denominator df=201 Power (1-β err prob)=0.3781274

^aPHQ: Patient Health Questionnaire

Figure 1. Flow of participants into the intervention group.

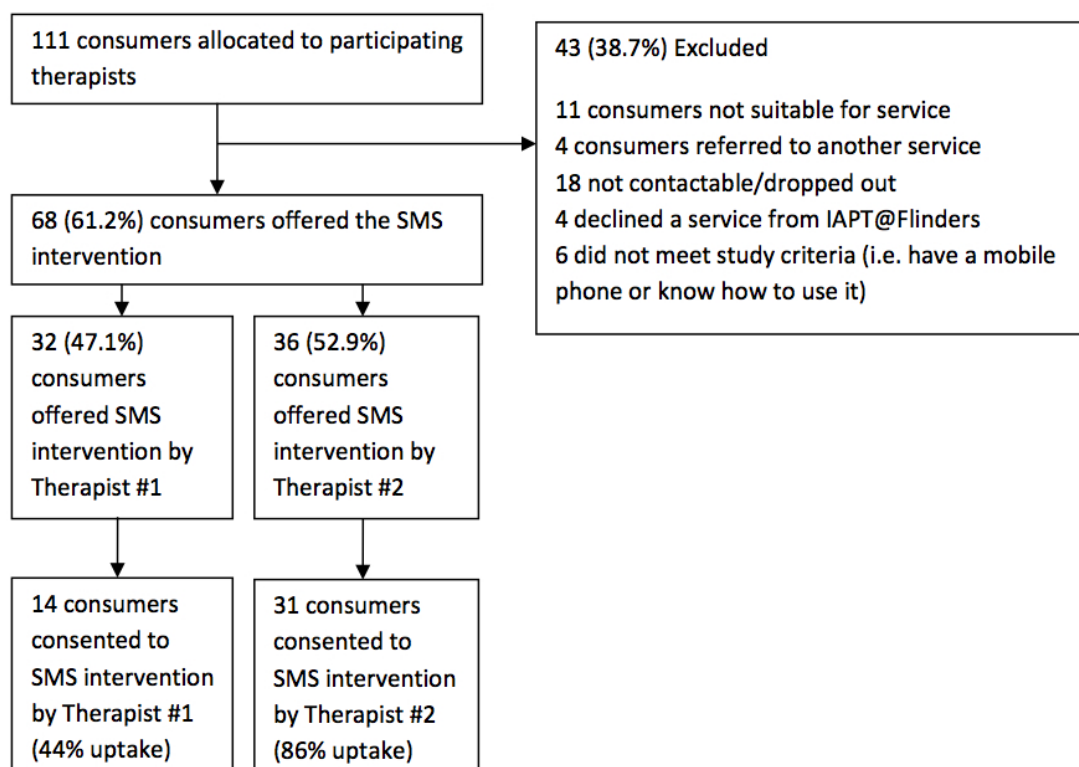


Figure 2. Flow of participants into the historical control group.

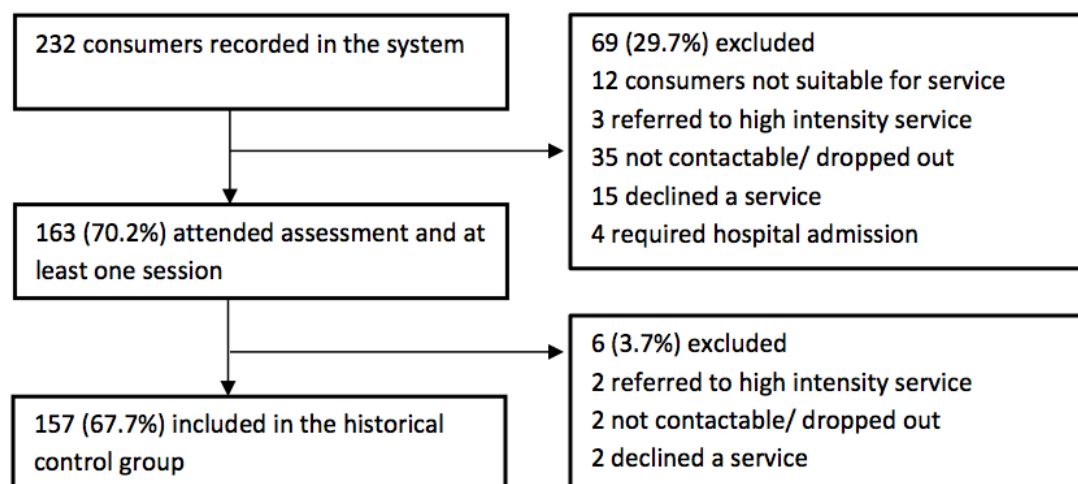
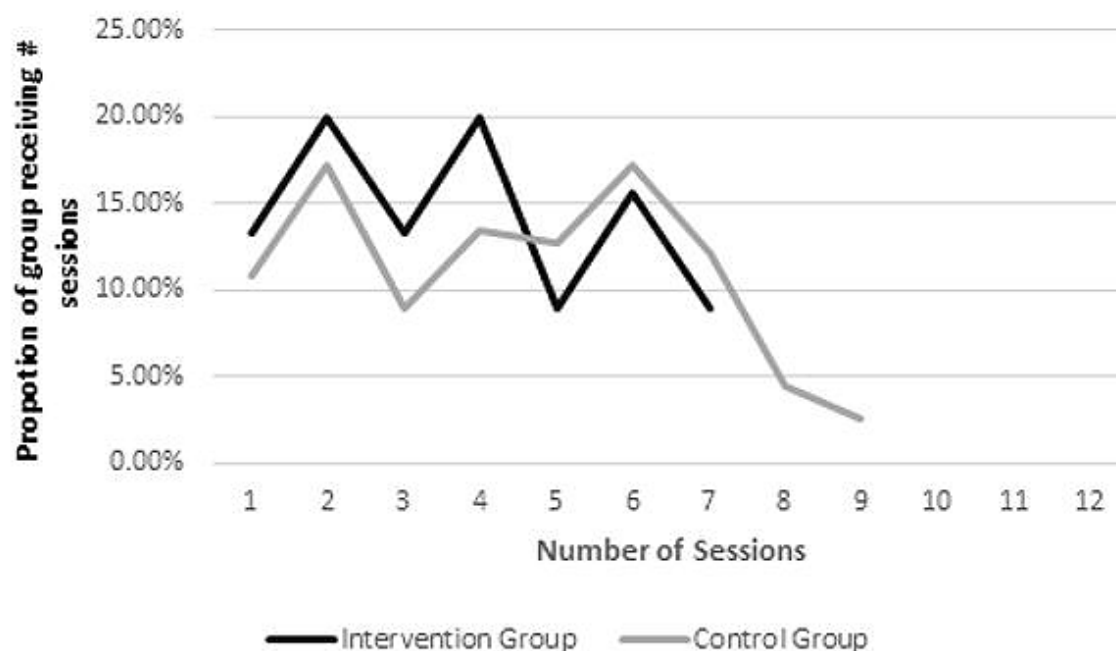


Figure 3. Proportion of intervention and control participants by number of sessions attended.



Discussion

Principal Findings

Given the paucity of controlled trials of supplementing psychotherapy with text messaging, we utilized a historical control group to explore the impact of adding individually tailored between-session text messages to telephone-based psychotherapy, delivered as part of a post-ED support service. Working directly with therapists, we developed a text messaging intervention that emphasized collaboration between therapist and consumer in composing and scheduling between-session text messages related directly to the content of sessions. With the assistance of an IT partner (goACT), the text messaging

intervention was implemented, with few practical issues, into an existing service.

To our knowledge, this is one of the first controlled studies of adding SMS, or text messaging, to psychotherapy. The study had high quality pre-post data because of the service’s session-by-session assessment protocol. The text messaging intervention was designed and implemented rapidly in practice, with informal feedback from clinicians and consumers indicating it was quick to learn, easy to implement, and a worthwhile addition to therapy. We anticipate the intervention could be transported easily to other clinical settings.

Intervention and control groups both showed significant improvements in depression, anxiety, and functional impairment

over time indicating positive impacts of the IAPT@Flinders program. In fact, using 'reliable change' criteria set out by Gyani and colleagues [36], we found 76% of our sample demonstrated significant positive improvements, compared with 63.7% reported in the first year of the UK IAPT program. As such, the standard care received by both groups had a strong positive effect on psychological symptoms.

Limitations

Consumers receiving the text messaging intervention in addition to standard care did not show significantly better clinical outcomes than a historical control group who received standard care without text messaging. A number of factors should be considered in interpreting this finding. First, we sought to detect the additional benefit of adding text messaging to what was already an efficacious treatment (low intensity telephone-based CBT) that was generating significant improvements in psychological symptoms. In contrast, other published trials have tested interventions against less active control groups that showed little change over time. Second, post-hoc power analyses suggested the study was underpowered to detect statistically significant effects of the size noted. Thus, genuine differences between the groups may have been obscured by lack of power. Third, the use of a historical control group, while a pragmatic solution for a counterfactual, raises some concerns about the comparability of the two groups and the causes of any observed differences. For example, the fewer number of sessions noted in the treatment group may be due to improvements in efficiency in the service over time.

With these limitations in mind, we provide the following suggestions for boosting the impact of the intervention and the power of the study design. The messages in the current study emphasized two core behavior change principles, namely, psychoeducation and homework/goal reminders. However, Michie and colleagues [37] have identified 93 behavior change techniques used in behavior change interventions; thus, many different behavior change techniques exist that were not utilized in the current study. The addition of other message types (eg, reward messages for goal completion, messages inviting consumers to monitor mood) may help supplement additional

psychotherapy processes. Second, text messages in the current trial were one-way only. Providing consumers with the opportunity to reply to messages (including rating their usefulness, notifying of goal completion) could increase consumer engagement with the messages and alert therapists to which messages are the most engaging. Finally, while SMS text messaging remains the dominant messaging platform, modern smartphones have access to multiple messaging platforms (eg, Whatsapp, Google hangouts, iMessage), which have additional features such as multimedia and emoticons that could enhance the richness of messages being sent to consumers.

Future evaluations of this type of intervention should aspire to be randomized controlled trials, although these can be difficult to implement in clinical practice due to limited resources, lack of personnel, risk of consumer disengagement, and contamination. An alternative is a design similar to that of the current study, but with some refinements: larger sample, structured follow-up, detailed attendance records, the use of multiple therapists, the collection of detailed intake information (including mobile phone ownership and use), and statistical techniques such as propensity score matching to improve matching of intervention and historical control participants. We also recommend the implementation of a session-by-session assessment protocol, which is both clinically useful [38] and also ensures at least two data points for the majority of individuals receiving a service.

Conclusions

In conclusion, the addition of individually tailored between-session text messaging to consumers receiving telephone-based psychotherapy following an ED admission, did not significantly improve clinical outcomes. However, this finding should be treated with caution given study design and power. The text messaging intervention was easy to implement and received positive feedback from therapists and consumers. For future trials, we specifically recommend refinement of the text messaging protocol as well as a randomized controlled trial study design to investigate whether the addition of mobile phone-based intervention components to psychotherapy can enhance consumer outcomes.

Acknowledgments

IAPT@Flinders would like to thank the team at goACT (goact.com.au) for providing access to their Web-based client management software that managed the sending of text messages in this trial.

Conflicts of Interest

NB is a shareholder and founding member of goACT.

References

1. Clough BA, Casey LM. Technological adjuncts to increase adherence to therapy: a review. *Clin Psychol Rev* 2011 Jul;31(5):697-710. [doi: [10.1016/j.cpr.2011.03.006](https://doi.org/10.1016/j.cpr.2011.03.006)] [Medline: [21497153](https://pubmed.ncbi.nlm.nih.gov/21497153/)]
2. Proudfoot J. The future is in our hands: the role of mobile phones in the prevention and management of mental disorders. *Aust N Z J Psychiatry* 2013 Feb;47(2):111-113. [doi: [10.1177/0004867412471441](https://doi.org/10.1177/0004867412471441)] [Medline: [23382507](https://pubmed.ncbi.nlm.nih.gov/23382507/)]
3. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. *Professional Psychology: Research and Practice* 2011;42(6):505-512. [doi: [10.1037/a0024485](https://doi.org/10.1037/a0024485)]

4. Boschen MJ, Casey LM. The use of mobile telephones as adjuncts to cognitive behavioral psychotherapy. *Professional Psychology: Research and Practice* 2008;39(5):546-552. [doi: [10.1037/0735-7028.39.5.546](https://doi.org/10.1037/0735-7028.39.5.546)]
5. Clough BA, Casey LM. Technological adjuncts to enhance current psychotherapy practices: a review. *Clin Psychol Rev* 2011 Apr;31(3):279-292. [doi: [10.1016/j.cpr.2010.12.008](https://doi.org/10.1016/j.cpr.2010.12.008)] [Medline: [21382535](https://pubmed.ncbi.nlm.nih.gov/21382535/)]
6. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. *J Ment Health* 2011 Dec;20(6):509-524. [doi: [10.3109/09638237.2011.608746](https://doi.org/10.3109/09638237.2011.608746)] [Medline: [21988230](https://pubmed.ncbi.nlm.nih.gov/21988230/)]
7. Preziosa A, Grassi A, Gaggioli A, Riva G. Therapeutic applications of the mobile phone. *British Journal of Guidance & Counselling* 2009 Aug;37(3):313-325. [doi: [10.1080/03069880902957031](https://doi.org/10.1080/03069880902957031)]
8. Agyapong VIO, Farren CK, McLoughlin DM. Mobile phone text message interventions in psychiatry - what are the possibilities? *CPSR* 2011 Feb 01;7(1):50-56. [doi: [10.2174/157340011795945847](https://doi.org/10.2174/157340011795945847)]
9. Spaniel F, Vohlídková P, Hrdlická J, Kožený J, Novák T, Motlová L, et al. ITAREPS: information technology aided relapse prevention programme in schizophrenia. *Schizophr Res* 2008 Jan;98(1-3):312-317. [doi: [10.1016/j.schres.2007.09.005](https://doi.org/10.1016/j.schres.2007.09.005)] [Medline: [17920245](https://pubmed.ncbi.nlm.nih.gov/17920245/)]
10. Kramer I, Simons CJ, Hartmann JA, Menne-Lothmann C, Viechtbauer W, Peeters F, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. *World Psychiatry* 2014 Feb;13(1):68-77 [FREE Full text] [doi: [10.1002/wps.20090](https://doi.org/10.1002/wps.20090)] [Medline: [24497255](https://pubmed.ncbi.nlm.nih.gov/24497255/)]
11. Kauer SD, Reid SC, Croke AHD, Khor A, Hearps SJC, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. *J Med Internet Res* 2012;14(3):e67 [FREE Full text] [doi: [10.2196/jmir.1858](https://doi.org/10.2196/jmir.1858)] [Medline: [22732135](https://pubmed.ncbi.nlm.nih.gov/22732135/)]
12. Reid S, Kauer S, Hearps S, Croke A, Khor A, Sancu L, et al. A mobile phone application for the assessment and management of youth mental health problems in primary care: a randomised controlled trial. *BMC Fam Pract* 2011;12:131 [FREE Full text] [doi: [10.1186/1471-2296-12-131](https://doi.org/10.1186/1471-2296-12-131)] [Medline: [22123031](https://pubmed.ncbi.nlm.nih.gov/22123031/)]
13. Pijnenborg GH, Withaar FK, Brouwer WH, Timmerman ME, van den Bosch RJ, Evans JJ. The efficacy of SMS text messages to compensate for the effects of cognitive impairments in schizophrenia. *Br J Clin Psychol* 2010 Jun;49(Pt 2):259-274. [doi: [10.1348/014466509X467828](https://doi.org/10.1348/014466509X467828)] [Medline: [19735607](https://pubmed.ncbi.nlm.nih.gov/19735607/)]
14. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. *Psychiatry Res* 2012 Dec 30;200(2-3):89-95. [doi: [10.1016/j.psychres.2012.07.034](https://doi.org/10.1016/j.psychres.2012.07.034)] [Medline: [22901437](https://pubmed.ncbi.nlm.nih.gov/22901437/)]
15. Bauer S, Okon E, Meermann R, Kordy H. Technology-enhanced maintenance of treatment gains in eating disorders: efficacy of an intervention delivered via text messaging. *J Consult Clin Psychol* 2012 Aug;80(4):700-706. [doi: [10.1037/a0028030](https://doi.org/10.1037/a0028030)] [Medline: [22545736](https://pubmed.ncbi.nlm.nih.gov/22545736/)]
16. Agyapong VI, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. *J Affect Disord* 2012 Dec 10;141(2-3):168-176. [doi: [10.1016/j.jad.2012.02.040](https://doi.org/10.1016/j.jad.2012.02.040)] [Medline: [22464008](https://pubmed.ncbi.nlm.nih.gov/22464008/)]
17. Marasinghe RB, Edirippulige S, Kavanagh D, Smith A, Jiffry M. Effect of mobile phone-based psychotherapy in suicide prevention: a randomized controlled trial in Sri Lanka. *J Telemed Telecare* 2012 Apr;18(3):151-155. [doi: [10.1258/jtt.2012.SFT107](https://doi.org/10.1258/jtt.2012.SFT107)] [Medline: [22362830](https://pubmed.ncbi.nlm.nih.gov/22362830/)]
18. Baer L, Minichiello WE, Jenike MA, Holland A. Use of a portable computer program to assist behavioral treatment in a case of obsessive compulsive disorder. *J Behav Ther Exp Psychiatry* 1988 Sep;19(3):237-240. [Medline: [3235693](https://pubmed.ncbi.nlm.nih.gov/3235693/)]
19. Kimhy D, Corcoran C. Use of Palm computer as an adjunct to cognitive-behavioural therapy with an ultra-high-risk patient: a case report. *Early Interv Psychiatry* 2008 Nov;2(4):234-241 [FREE Full text] [doi: [10.1111/j.1751-7893.2008.00083.x](https://doi.org/10.1111/j.1751-7893.2008.00083.x)] [Medline: [19884956](https://pubmed.ncbi.nlm.nih.gov/19884956/)]
20. Przeworski A, Newman MG. Palmtop computer-assisted group therapy for social phobia. *J Clin Psychol* 2004 Feb;60(2):179-188. [doi: [10.1002/jclp.10246](https://doi.org/10.1002/jclp.10246)] [Medline: [14724925](https://pubmed.ncbi.nlm.nih.gov/14724925/)]
21. Norton M, Wonderlich SA, Myers T, Mitchell JE, Crosby RD. The use of palmtop computers in the treatment of bulimia nervosa. *Eur Eat Disorders Rev* 2003 May;11(3):231-242. [doi: [10.1002/erv.518](https://doi.org/10.1002/erv.518)]
22. Newman MG, Consoli AJ, Taylor CB. A palmtop computer program for the treatment of generalized anxiety disorder. *Behav Modif* 1999 Oct;23(4):597-619. [Medline: [10533442](https://pubmed.ncbi.nlm.nih.gov/10533442/)]
23. Aguilera A, Muñoz RF. Text messaging as an adjunct to CBT in low-income populations: a usability and feasibility pilot study. *Professional Psychology: Research and Practice* 2011;42(6):472-478. [doi: [10.1037/a0025499](https://doi.org/10.1037/a0025499)]
24. Rizvi SL, Dimeff LA, Skutch J, Carroll D, Linehan MM. A pilot study of the DBT coach: an interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. *Behav Ther* 2011 Dec;42(4):589-600. [doi: [10.1016/j.beth.2011.01.003](https://doi.org/10.1016/j.beth.2011.01.003)] [Medline: [22035988](https://pubmed.ncbi.nlm.nih.gov/22035988/)]
25. Muench F, van Stolk-Cooke K, Morgenstern J, Kuerbis A, Markle K. Understanding messaging preferences to inform development of mobile goal-directed behavioral interventions. *J Med Internet Res* 2014;16(2):e14 [FREE Full text] [doi: [10.2196/jmir.2945](https://doi.org/10.2196/jmir.2945)] [Medline: [24500775](https://pubmed.ncbi.nlm.nih.gov/24500775/)]

26. Owens C, Farrand P, Darvill R, Emmens T, Hewis E, Aitken P. Involving service users in intervention design: a participatory approach to developing a text-messaging intervention to reduce repetition of self-harm. *Health Expect* 2011 Sep;14(3):285-295. [doi: [10.1111/j.1369-7625.2010.00623.x](https://doi.org/10.1111/j.1369-7625.2010.00623.x)] [Medline: [20860777](https://pubmed.ncbi.nlm.nih.gov/20860777/)]
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](https://doi.org/10.1016/j.amepre.2008.09.040)] [Medline: [19135907](https://pubmed.ncbi.nlm.nih.gov/19135907/)]
28. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. *Epidemiol Rev* 2010 Apr;32(1):56-69 [FREE Full text] [doi: [10.1093/epirev/mxq004](https://doi.org/10.1093/epirev/mxq004)] [Medline: [20354039](https://pubmed.ncbi.nlm.nih.gov/20354039/)]
29. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. *J Telemed Telecare* 2011;17(1):41-48. [doi: [10.1258/jtt.2010.100322](https://doi.org/10.1258/jtt.2010.100322)] [Medline: [21097565](https://pubmed.ncbi.nlm.nih.gov/21097565/)]
30. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. *Int Rev Psychiatry* 2011 Aug;23(4):318-327 [FREE Full text] [doi: [10.3109/09540261.2011.606803](https://doi.org/10.3109/09540261.2011.606803)] [Medline: [22026487](https://pubmed.ncbi.nlm.nih.gov/22026487/)]
31. Lovell K, Richards D. A recovery programme for depression. London: Rethink; 2008.
32. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
33. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
34. Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002 May;180:461-464 [FREE Full text] [Medline: [11983645](https://pubmed.ncbi.nlm.nih.gov/11983645/)]
35. 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](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
36. Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: lessons from year one of IAPT. *Behav Res Ther* 2013 Sep;51(9):597-606 [FREE Full text] [doi: [10.1016/j.brat.2013.06.004](https://doi.org/10.1016/j.brat.2013.06.004)] [Medline: [23872702](https://pubmed.ncbi.nlm.nih.gov/23872702/)]
37. 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](https://doi.org/10.1007/s12160-013-9486-6)]
38. Reese RJ, Norsworthy LA, Rowlands SR. Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy (Chic)* 2009 Dec;46(4):418-431. [doi: [10.1037/a0017901](https://doi.org/10.1037/a0017901)] [Medline: [22121836](https://pubmed.ncbi.nlm.nih.gov/22121836/)]

Abbreviations

ANOVA: analysis of variance
CBT: cognitive behavioral therapy
DNA: did not attend
ED: emergency department
GAD-7: Generalized Anxiety Disorder questionnaire
IAPT: Improving Access to Psychological Therapies
PHQ-9: Patient Health Questionnaire
SMS: short message service
WSAS: Work and Social Adjustment Scale

Edited by G Eysenbach; submitted 15.11.13; peer-reviewed by A Aguilera, J Torous, S Kauer; comments to author 19.01.14; revised version received 06.04.14; accepted 07.07.14; published 08.10.14.

Please cite as:

Furber G, Jones GM, Healey D, Bidargaddi N

A Comparison Between Phone-Based Psychotherapy With and Without Text Messaging Support In Between Sessions for Crisis Patients
J Med Internet Res 2014;16(10):e219

URL: <http://www.jmir.org/2014/10/e219/>

doi: [10.2196/jmir.3096](https://doi.org/10.2196/jmir.3096)

PMID: [25295667](https://pubmed.ncbi.nlm.nih.gov/25295667/)

©Gareth Furber, Gabrielle Margaret Jones, David Healey, Niranjana Bidargaddi. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 08.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly

cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Use of Mobile Phone Text Message Reminders in Health Care Services: A Narrative Literature Review

Kati Anneli Kannisto^{1,2}, BNSc; Marita Hannele Koivunen^{1,2}, PhD; Maritta Anneli Välimäki^{1,3}, RN, PhD

¹Department of Nursing Science, University of Turku, Turku, Finland

²Satakunta Hospital District, Pori, Finland

³Turku University Hospital, Turku, Finland

Corresponding Author:

Kati Anneli Kannisto, BNSc
Department of Nursing Science
University of Turku
Lemminkäisenkatu 1
Turku, 20014
Finland
Phone: 358 50 5312652
Fax: 358 2 3338400
Email: kaankan@utu.fi

Abstract

Background: Mobile text messages are a widely recognized communication method in societies, as the global penetration of the technology approaches 100% worldwide. Systematic knowledge is still lacking on how the mobile telephone text messaging (short message service, SMS) has been used in health care services.

Objective: This study aims to review the literature on the use of mobile phone text message reminders in health care.

Methods: We conducted a systematic literature review of studies on mobile telephone text message reminders. The data sources used were PubMed (MEDLINE), CINAHL, Proquest Databases/ PsycINFO, EMBASE, Cochrane Library, Scopus, and hand searching since 2003. Studies reporting the use of SMS intended to remind patients in health services were included. Given the heterogeneity in the studies, descriptive characteristics, purpose of the study, response rates, description of the intervention, dose and timing, instruments, outcome measures, and outcome data from the studies were synthesized using a narrative approach.

Results: From 911 initial citations, 60 studies were included in the review. The studies reported a variety of use for SMS. Mobile telephone text message reminders were used as the only intervention in 73% (44/60) of the studies, and in 27% (16/60) of the remaining studies, SMS was connected to another comprehensive health intervention system. SMS reminders were sent to different patient groups: patients with HIV/AIDS (15%, 9/60) and diabetes (13%, 8/60) being the most common groups. The response rates of the studies varied from 22-100%. Typically, the text message reminders were sent daily. The time before the specific intervention to be rendered varied from 10 minutes (eg, medication taken) to 2 weeks (eg, scheduled appointment). A wide range of different evaluation methods and outcomes were used to assess the impact of SMS varying from existing databases (eg, attendance rate based on medical records), questionnaires, and physiological measures. About three quarters of the studies (77%, 46/60) reported improved outcomes: adherence to medication or to treatment reportedly improved in 40% (24/60) of the studies, appointment attendance in 18% (11/60) of the studies, and non-attendance rates decreased in 18% (11/60) of the studies. Other positive impacts were decreased amount of missed medication doses, more positive attitudes towards medication, and reductions in treatment interruptions.

Conclusions: We can conclude that although SMS reminders are used with different patient groups in health care, SMS is less systematically studied with randomized controlled trial study design. Although the amount of evidence for SMS application recommendations is still limited, having 77% (46/60) of the studies showing improved outcomes may indicate its use in health care settings. However, more well-conducted SMS studies are still needed.

(*J Med Internet Res* 2014;16(10):e222) doi:[10.2196/jmir.3442](https://doi.org/10.2196/jmir.3442)

KEYWORDS

text messaging; short message service; cellular phone; reminder system; review

Introduction

With more than 6.8 billion mobile phone users and mobile phone technology penetration near 100% worldwide, mobile technology and text messages have changed communication between people [1] and increased the use of this technology in health care services [2]. Mobile phones are used in low-income countries [3,4] and in most social groups [2] including patients with psychiatric problems [5]. Due to its low costs, quick delivery [2], safety issues [6], and reduced intrusiveness compared to phone calls [5], mobile technology has been favored in various contexts and is recommended in a variety of strategies [7,8] and guidelines [9,10]. However, implementing new interventions requires continuous education and training among staff members [11].

The use of text messaging (short message service, SMS) applications for behavioral change is at an early stage of research [3]. Systematic reviews have already been conducted in this area, although discrepancies between the results of the previous reviews can be found. Previous reviews have shown that SMS reminders had a positive impact on patient appointment attendance [12], adherence to chronic medication [13] and to antiretroviral therapy [14], patient self-management [15] or health outcomes, and care processes [16]. On the contrary, Gurol-Urganci et al [17] found very limited evidence that communicating results of medical investigations by SMS would be useful. Kauppi et al [18] as well did not find clear evidence that information and communication technology (ICT)-based prompting (like SMS) would improve medication adherence with people with serious mental illness. However, little is known about which specific patient groups SMS reminders have been used for in health care. To form a more coherent picture of how SMS reminders have been used in clinical practice and to provide a more thorough understanding of the knowledge

accumulated in the area, it is important to figure out the context, situations, and audience for past text message reminder use and the possible benefits to patients. Therefore, this review aims to synthesize studies investigating the use of mobile phone text message reminders in health care. The review was guided by the following questions: (1) What purposes have text message reminders been used for in health care?, (2) How have the impacts of text message reminders been assessed?, and (3) What are the impacts of using text messages as reminders in health care?

Methods

Design

A systematic review design with narrative methods was used. More precisely, a review methodology [19] was conducted to form a conception of the use of mobile phone text messages as reminders in health care.

Search Strategy

We conducted a comprehensive literature search on February 21, 2013. The following electronic databases were searched with the help of an information specialist at the Medical Library: PubMed (MEDLINE), CINAHL, Proquest Databases/PsycINFO, Embase, Scopus, and the Cochrane Library. The search terms (or equivalent index terms and free-text words) for each of the databases were used to ensure a broad coverage of published studies in our review. Detailed search terms are presented in Table 1.

References were also collected by screening the reference lists of the 906 articles, and 2 more papers were found. In addition, a hand search in all JMIR journals was conducted (in August 2014) leading to 3 additional papers [20-22]. Thus, we identified a total of 911 published articles relevant to our topic.

Table 1. Databases and search terms used, and references found (N=906).

Database	Search terms	References, n
PubMed (MEDLINE)	("Cellular Phone"[Mesh] OR "cell phone"[tiab] OR "cell phones"[tiab] OR "cellular phone"[tiab] OR "mobile phone"[tiab] OR "mobile phones"[tiab] OR "short message service"[tiab] OR "short messaging service"[tiab] OR "text messaging"[tiab] OR "text messages"[tiab] OR "text message"[tiab] OR (sms[tiab] AND (message[tiab] OR messages[tiab] OR messaging[tiab]))) AND ("Reminder Systems"[Mesh] OR remind*[tiab] OR prompt*[tiab])	315
CINAHL	(MH "Wireless Communications" OR MH "Telephone" OR MH "Instant Messaging" OR TI ("short message service" OR "short messaging service" OR "text messaging" OR "text messages" OR "text message" OR (sms AND (message OR messages OR messaging)))) OR AB ("short message service" OR "short messaging service" OR "text messaging" OR "text messages" OR "text message" OR (sms AND (message OR messages OR messaging)))) AND (MH "Reminder Systems" OR TI remind* OR TI prompt* OR AB remind* OR AB prompt*)	194
Proquest Databases/ PsycINFO	all(reminder*) AND all(("cellular phone*" OR sms OR "short text message*" OR "text message*" OR "cell phone*" OR "mobile phone*")	75
Embase	((sms OR 'short text message' OR 'short text messages' OR 'text messages' OR 'text message' OR 'text messaging'/exp OR 'text messaging' OR 'mobile phone'/exp OR 'mobile phone' OR 'mobile phones' OR 'cell phone'/exp OR 'cell phone' OR 'cell phones' OR 'cellular phone'/exp OR 'cellular phone' OR 'cellular phones') AND ('reminder system'/exp OR 'reminder system')) OR 'sms reminder' OR 'sms reminders'	179
Scopus	((((sms OR "short text message*" OR "text message*" OR "cell phone*" OR "mobile phone*" OR "cellular phone*") AND ("reminder system*")) OR "sms reminder*") AND (health	143

Eligibility Criteria

The review was limited to texts published in English, with an abstract available, between 2003 and 2013. The limitation in publication years was chosen due to a marked increase in information technology during the last decade [23]. The review was also limited to studies of text message reminders in the health care domain, sent from health care services to patients' mobile phones. Patients of all ages and with any diagnoses were included. Further, we included only peer-reviewed, published papers using a variety of design and research methods.

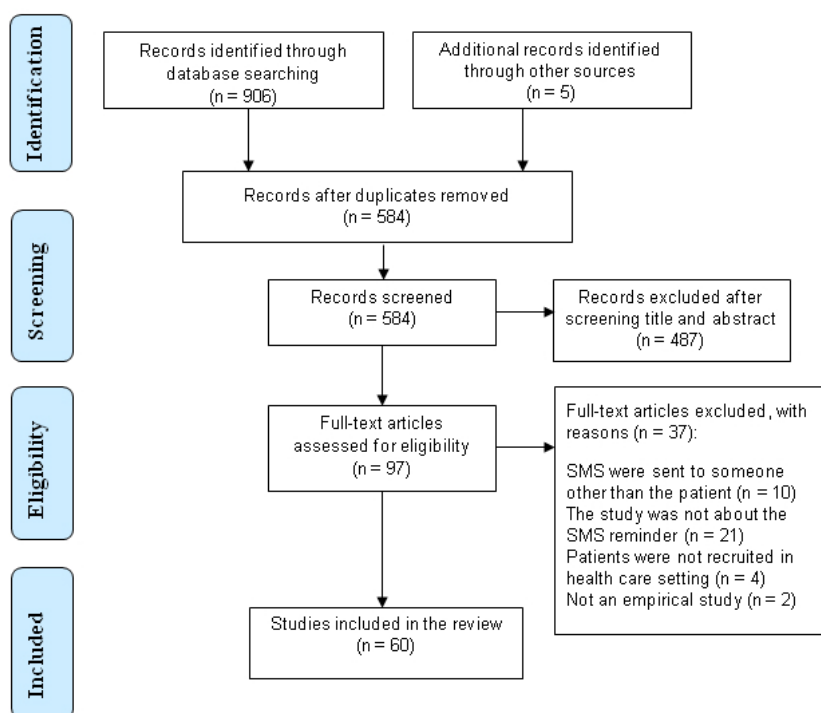
Studies were excluded if SMSs were received by a parent, relative or friend, health care student, or staff member; if a reminder was sent by email or letter; or if SMS was used for non-clinical purposes (eg, for the recruitment of study participants, to survey patients' willingness to receive text messages). Further, papers describing the design process of the SMS system, theoretical papers, statistical reviews, books or

book chapters, letters, dissertations, editorials, and study protocols were excluded.

Study Selection

The study selection consisted of four steps. First, 2 authors (KK, MK) independently screened all titles and abstracts ($n=911$) of relevance for this systematic review [24]. Second, the abstracts of all relevant articles were screened for eligibility by the same 2 authors. Third, the full papers of the included publications were obtained and screened (KK) for inclusion and exclusion criteria. In case of any discrepancy between the decisions made, the papers were discussed until consensus was reached with the support of MV. Fourth, the reference lists of all papers included and systematic reviews identified in the original search were checked to find additional publications that met our inclusion criteria. After study selection, we had 60 studies to be extracted. Figure 1 outlines the search process of the literature [25].

Figure 1. PRISMA flow diagram outlining the review process.



Data Extraction

We created a specific data extraction grid to collect information systematically related to our aims in the synthesis study by one author (KK). The data extraction was based on the CONSORT-EHEALTH checklist [26] where possible.

Descriptive Characteristics of the Study

The following information of data items was collected [27]: (1) name(s) of the author(s), (2) year of publication (papers published since 2003 were included due to a marked increase in information technology during the last decade [23]), (3) country where the study was conducted, (4) purposes of the studies related to the use of mobile phone text message reminders, (5) setting of the studies was coded with a specific term used in the study (eg, outpatient or inpatient clinics, general

or private hospitals), and the patient group who received text message reminders was extracted, (6) type (quantitative or qualitative) of the study, (7) study design, (8) number of patients enrolled and participating in the study, and response rate of the studies were extracted, (9) intervention based on mobile phone text message reminders, (10) goals of text messages, (11) dose of the intervention based on mobile phone text message reminders (eg, the number of text message reminders, how often text message reminders were sent), and the timing (eg, the time of day or week the text message was sent; the time of a reminder before a specific intervention) of the intervention were extracted, (12) outcome measures as a key concept used, (13) instruments used to measure the outcomes of the intervention based on mobile telephone text message reminders (eg, names of the instruments used), and (14) outcomes of the intervention were extracted and described as increased, decreased, or unchanged.

Increase, decrease, or unchanged were then presented as arrows up (\uparrow), down (\downarrow) or horizontal (\Leftrightarrow) (respectively) (see [28]).

Analyses

The data on each included study were entered into the specific data extraction grid. Each study was treated as a separate case. Descriptive characteristics of the studies were categorized manually according to our research questions. The methodological quality of the studies was appraised with the Mixed Method Appraisal Tool (MMAT) by Pluye et al [29]. The method was designed to appraise the methodological quality of the studies in complex systematic literature reviews that include qualitative, quantitative, and mixed methods studies. For appraising qualitative studies, we used Section 1 of the MMAT, which contains items related to data sources, data analysis, context, and researcher's influence. Section 2 of the MMAT was used to appraise randomized controlled studies; it contains items related to randomization, allocation concealment, assessment of outcome data and completeness of follow-up (drop-out). Section 3 was used for non-randomized studies; it contains items related to participants' recruitment, outcome measurements, comparability of groups, and completeness of outcome data. Section 4 was used for descriptive studies; it contains items related to relevant sampling strategy, representativeness of the sample, outcome measurements, and acceptability of the response rate. Each item was scored as "yes", "no", or "can't tell" [29]. In 27% (16/60) of the studies, the quality score was 4/4, meaning that all four criteria were met. In 45% (27/60) of the studies, the quality score was 3/4. In 72% (43/60) of the included studies, the quality score was 3/4 or 4/4, indicating the methodological quality of the included studies.

Results

Study Selection

The literature search yielded 911 publications. Duplicates were removed leaving 584 papers for further abstract screening. Following that screening according to inclusion and exclusion criteria by the Centre for Reviews and Dissemination [24], we

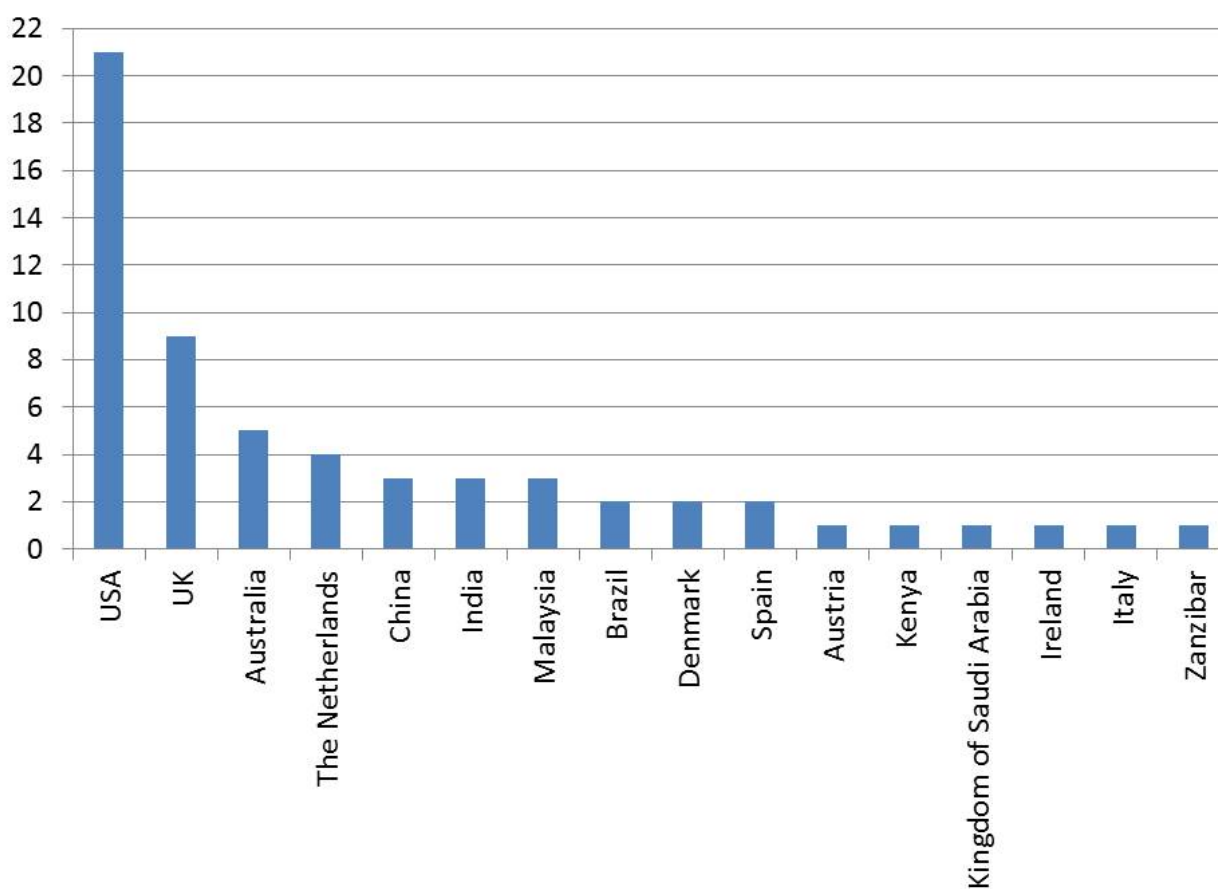
excluded 487 papers based on the title and the abstract. All together 97 potential articles were obtained for full-text review by 2 independent reviewers, of which 60 studies were included in the review for further data extraction.

Characteristics of the Included Studies

Author, year, country, setting, type of study, design, patient group, and sample were extracted to describe the characteristics of the studies. The authors of the studies are reported in each table dealing with the included studies (see [Multimedia Appendices 1-3](#)). The studies included in our analysis were published between 2004 and 2013. The number of published studies increased steadily until 2011, being highest in 2012. Of the included studies, 37% (22/60) studies were published in 2012. Except for one study [30], all publications involved outpatients. The studies were mostly conducted in the United States (35%, 21/60), followed by the United Kingdom and Australia ([Figure 2](#)).

Of the included studies, 95% (57/60) had a quantitative design, one (1) had a qualitative design, and two (2) used both quantitative and qualitative designs. Over one-third (35%, 21/60) of the studies were randomized controlled trials (RCT). Other studies were non-randomized feasibility studies, before-and-after studies, cross-sectional studies, retrospective and prospective studies, cohort studies with or without historical control, clinical trials, or qualitative descriptive studies.

The most common patient groups described were patients with human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS; 15%, 9/60), diabetes (13%, 8/60), asthma (8%, 5/60), or schizophrenia (7%, 4/60). Other patient groups are described in more detail in [Multimedia Appendix 2](#). The sample size of the studies varied from 4 participants to 9959. In half of the studies (53%, 32/60), the sample size was 100 or under, and in 23 studies (38%, 23/60), it was over 100. In 5 studies (8%, 5/60) the sample size was shown as the amount of appointments, not participants. The response rates of the studies varied from 22-100%. Descriptive characteristics of the included studies are presented in more detail in [Multimedia Appendices 1 and 2](#).

Figure 2. Countries of the publications included.

Purposes of Text Message Reminders Used in Health Care

The purpose of the study, description of the intervention, dose, and timing were extracted to describe the purposes of text message reminders used in health care. Mobile phone text message reminders were used to remind patients about their medication or treatment in 63% (38/60) of the studies, and both to increase the attendance to clinical appointments and to decrease the non-attendance to clinical appointments with patients from different patient groups in 37% (22/60) of the studies ([Multimedia Appendix 2](#)).

The mobile telephone text message reminders were used as the only intervention in 73% (44/60) of the studies ([Multimedia Appendix 2](#)). In 27% (16/60) of the studies, the approach was multifaceted, indicating that text message reminders were connected to another comprehensive health intervention system, such as educational text messages (7%, 4/60) [31-34], informational text messages about patients' disease and healthy living (12%, 7/60) [20,21,32,33,35-37], supportive text messages [22,38] or to diary data collection [39]. In the studies by Rodrigues et al [40] and Sidney et al [41], reminders were sent as non-interactive neutral pictures delivered as an SMS.

The dose and timing of the text message reminders depended on the dosage of the medication or treatment and a patient's scheduled appointment ([Multimedia Appendix 2](#)). In 55%

(33/60) of the studies, the dose was reported as how often the text message reminder was sent, so that the most common dose was to send the text message reminders daily (35%, 21/60) [20-22,31,32,35,38,39,42-54]. In 38% (23/60) of the studies, the dose was reported as amount of text message reminders sent, and in 7% (4/60) of the studies, the dose was reported to depend on patients' preferences.

The timing was reported as the time of the day (eg, morning, evening) or as a certain time (eg, 10 a.m., 8 p.m.) in 37% (22/60) of the included studies, and as how many days before the appointment the reminder was sent (eg, one day before the appointment) in 25% (15/60) of the studies. In 22% (13/60) of the studies, the timing was reported to be based on patients' personal needs. However, in 17% (10/60) of the studies, the time of sending text messages was not reported. Further, the timing varied from 10 minutes before the medication was due to be taken [55] to 2 weeks before the scheduled appointment [56]. The dose and timing of the text message reminders are shown more specifically in [Multimedia Appendix 2](#).

Assessment Methods to Evaluate the Impact of Mobile Phone Text Message Reminders

A description of the instruments used as an assessment method was extracted ([Multimedia Appendix 3](#)). In 43% (26/60) of the studies, the impact was assessed using existing databases (eg, attendance rate/did not attend rate) [34,39,52,56-78].

In 12% (7/60) of the studies, the impact of the mobile phone text message reminders was assessed using questionnaires [20,32,41,49-51,53], and in 18% (11/60) of the studies, the impact was assessed using physiological measures [22,31,35,37,44,45,47,48,52,79,80]. Out of these 11 physiological assessments, patients' self-assessment was used alone in one study, blood test alone in one study, self-reported weight in one study, and in eight studies patients' self-assessment was connected with electronic monitoring, a questionnaire, pill counting, or a blood test. Other assessment methods were electronic monitoring alone (n=3) [42,46,81], pill count alone (n=1) [40], system usage calculation (n=1) [36],

proportion of days covered calculation (n=1) [82], interview (n=1) [54], and observational measurement (n=2) [30,83].

In 10% (6/60) of the studies, the impact was assessed through patient satisfaction with the text message-based intervention [33,38,55,84-86]. In addition to these six studies, patient satisfaction was assessed in 15 studies. Patients' satisfaction with the text message reminders was assessed in total in 20 studies, and patients' reminder preferences in one study (Table 2). Patient satisfaction was assessed by questionnaires (n=14) [20-22,31-33,38,41,48,53,55,70,79,85], and by interviews (n=6) [36,39,44,73,80,84]. Patients' preferences regarding reminders (n=1) were assessed by calculating the percentage of patients who selected the SMS reminders [86].

Table 2. Assessment of patient satisfaction.

Author (year)	Outcome measure	Instruments	Outcomes ^a
Anhøj & Møldrup (2004) [39]	Feasibility of using SMS for asthma diary data collection	Focus group interview	+
	Participants' experiences with medication adherence reminders	Focus group interview	+
Agyapong et al (2013) [38]	Usefulness	Semistructured questionnaire	+
	Patient satisfaction with abstinence reminders	Semistructured questionnaire	+
	Patient satisfaction with medication reminders	Semistructured questionnaire	-
Arora et al (2012) [32]	Satisfaction with the TEXT-MED program	Questionnaire	+
Branson et al (2011) [70]	Patient satisfaction with text message reminders	Questionnaire	+
Britto et al (2011) [85]	Usefulness	Questionnaire	+
	Acceptability	Questionnaire	+
da Costa et al (2012) [80]	Patient satisfaction	Interview	+
Dick et al (2011) [44]	Satisfaction with the text message-based program	Interview	+
Dowshen et al (2012) [48]	Feasibility	"Satisfaction survey"	+
	Acceptability	"Satisfaction survey"	+
Fischer et al (2012) [73]	Feasibility	Focus group interview	+
Furberg et al (2012) [21]	Patient satisfaction with text messages	Satisfaction survey via SMS	+
Greaney et al (2012) [86]	Automated reminder preferences	SMS calculation	28%
Hanauer et al (2009) [36]	Feasibility	Interview	+
Holtz & Whitten (2009) [84]	Feasibility	Interview	+
	Compliance with monitoring asthma	Log-in records	+
Kollman et al (2007) [79]	Feasibility and user acceptance	Questionnaire	+
Lewis et al (2013) [53]	Receptivity to adherence messaging	Message receptivity questions via two-way text messages	+
	Clinical outcomes	Blood test (total virus load and CD4 counts)	+
Lua et al (2012) [33]	Feasibility and acceptability	Feedback form	+
Mao et al (2008) [55]	Patient satisfaction	Standardized questionnaire	+
Nundy et al (2013) [20]	Feasibility and acceptability	Patient experience survey	+
Pena-Robichaux et al (2010) [31]	Usability and satisfaction of the TM system	Questionnaire	+
Shaw et al (2013) [22]	Feasibility and acceptability	Questionnaire	+
Sidney et al (2012) [41]	Usefulness	Structured questionnaire	+
	Reminder preference:	Structured questionnaire, data on the delivery	
		Voice reminder	
	SMS alone		11%

^a+ patients' positive feedback, - patients' negative feedback.

Impact of Using Text Messages as Reminders in Health Care

Outcome measures and outcomes were extracted to describe the impacts of using text messages as reminders in health care.

Of the included studies, the outcome measures were adherence to medication or treatment (50%, 30/60), appointment attendance (22%, 13/60), appointment non-attendance (18%, 11/60), or patient satisfaction (10%, 6/60). The impacts of using SMS text

messages as reminders in health care are described in [Table 2](#) and [Multimedia Appendix 3](#).

Out of 60 studies, the outcomes reportedly improved in 77% (46/60). First, adherence to medication or to treatment improved in 24 studies [[20,21,31,32,35-37,39,40,43-46,48-51,53,54,66,79-82](#)]. Second, appointment attendance was reported to have improved in 11 studies [[30,34,57,60,61,69,70,74,75,77,78](#)]. Third, non-attendance rates reportedly decreased in 11 studies [[56,59,62-65,67,68,71,72,76](#)]. In addition, patients' attitudes towards medication were reported to have improved [[51](#)], the number of missed medication doses reportedly decreased [[44,81](#)], and text messages were found to have reduced treatment interruptions [[46](#)].

Outcomes in patient satisfaction were positive in those studies (n=6) where no impact was assessed [[33,38,55,84-86](#)]. In addition to these studies, patient satisfaction was assessed in 15 studies together with the impact assessment. In patients' opinions, text messages were easy to use [[20,44,48](#)], they reminded patients to take their medication [[32](#)], patients were willing to receive text messages [[53,66](#)], and they were satisfied with text messages [[69](#)].

Using text messages had advantages over other reminding systems. Text messages could be sent to patients simultaneously, they were always available [[59](#)], cost-effective [[59,63](#)], and sending text messages to patients required less staff [[63](#)]. Liew et al [[65](#)] found that text messages were as effective as telephone reminders but were low-priced [[60,61](#)]. However, in the study by Greaney et al [[86](#)], participants preferred automated voice response reminders (72%) instead of SMS reminders (28%). Patients' opinions about the usefulness of the text messages received varied from 88% [[31](#)] to 66% [[81](#)].

However, daily text message reminders did not improve adherence to oral contraceptive pills [[42](#)], acne treatment [[47](#)], or lupus erythematosus treatment [[52](#)]. Pijnenborg et al [[30,83](#)] found that the overall effect of prompting disappeared after the text message reminders ceased, indicating the dependence on continuous use of the intervention. Bos et al [[58](#)] and Fischer et al [[73](#)] found that there were no differences in appointment attendance before and after sending text message reminders.

Despite all the benefits and beneficial characteristics of the mobile phone text message reminders, this literature review shows that there are limitations to using mobile phone text message reminders. First, patients had privacy concerns about losing their mobile phones and other people possibly gaining access to the messages [[67,70](#)]. Although most (93%, 56/60) of the studies reported that the messages did not include the patient's name or other identification in the reminder message, four exceptions were found [[47,60,61,72](#)]. No adverse events were reported. Second, patients may have changed their mobile phone numbers without informing the health care staff [[60,61,67](#)], thus the staff could not be sure that all participants had received the text messages. Downing et al [[77](#)] found that the proportion of undelivered text messages was high. Koshy et al [[63](#)] demonstrated that patients may have not received the text message reminders due to incorrect data entry. Third, it is possible that patients adapted to the messages and stopped reading them [[43](#)].

Discussion

Principal Findings

The results of this narrative literature review showed that mobile phones and text messages are used worldwide, which supports the global penetration of mobile phone subscriptions [[1](#)] in different user groups in health care [[2,12,14,15](#)]. We have demonstrated that mobile phone text messages may have their uses in reminding patients about medication adherence [[43,45,47,50](#)] and in reducing non-attendance rates [[59,61,63,76](#)]. The possibility of using text message reminders as the only intervention or in conjunction with some other comprehensive health intervention systems further adds to the usability of text messages in health care services. Thus, SMS reminders deserve more attention as a potential innovation to improve health care operations [[87](#)].

On the other hand, some concerns were also identified. First, in 4 studies, the dose of text message reminders (eg, the number of text message reminders, how often text message reminders were sent), and in 12 studies, the timing of the text message reminders reported were based on patients' personal needs. Second, despite the safety of the text messages in health care [[6](#)], the literature review demonstrated privacy concerns, such as loss of a mobile phone or other people reading the messages [[67,70](#)]. Therefore, more emphasis should be put on how to guarantee that health-related patient information in electronic systems is anonymous and neutral enough to be managed even in open electronic systems. Special considerations are also required in designing the content of the reminder messages, entering the patient data to the automatic systems or dialing correct mobile phone numbers to protect patients' privacy and security issues [[4,88](#)].

The impacts of text message reminders focused on improving adherence to medication and improving appointment attendance. Although no meta-analysis was used due to the high heterogeneity of the data gathered, this review demonstrated that text message reminders were easy to use, useful for patients, they were willing to receive text messages, and satisfied with the text message reminders. This knowledge is essential because patients' views influence the acceptance of the text message intervention and its integration into patients' daily lives [[13](#)]. On the other hand, patients may adapt to the messages and the effectiveness of the messages diminishes. This is what happened in the study by Strandbygaard et al [[43](#)]; participants in the intervention group stopped reading their reminder messages after a few weeks.

Limitations

We recognize that there are some limitations in our review. First, the literature search yielded studies that were diverse methodologically and clinically. As the studies included were heterogeneous in study design, patient group studied, sample size, description of the intervention, and outcome measures, we synthesized the data with a narrative method, rather than trying to do a meta-analysis. As such, our findings cannot be used to recommend any preferred strategy for the use of mobile phone text message reminders in health care. Second, the studies differed in their methodological quality, which may have had

an impact on the results, and biased our findings and limited our interpretations. Third, we included studies only from peer-reviewed English-language journals, which may have restricted our findings and biased the data toward positive results. And fourth, we excluded studies if text messages were received by parents, relatives, or friends, which may exclude a wide variety of studies (eg, immunization reminders; see [89,90]) in different fields and further affected the conclusions from the review.

Future Research

Further evaluation of mobile phone text message reminder interventions is needed to form a more coherent picture of their use and effectiveness in health care services. This should be done with rigorous RCT studies of their effectiveness and cost-effectiveness. The research should also focus more on service users' and their caregivers' needs and preferences regarding the text message reminders to be received and how to maintain interest in text message reminders to achieve the

best possible impact. In addition, the assessment of users' satisfaction toward intervention should also be ensured. In this task, qualitative evaluations could also be used to hear users' voices. More research is also needed to ascertain the best ways to guarantee privacy and security in mobile phone text message reminder interventions.

Conclusions

The findings of this literature review are encouraging. However, the amount of evidence for SMS application recommendations is still limited. In our review, having 77% (46/60) of the studies showing improved outcomes may still indicate its use in health care settings. Although no firm conclusions can be drawn so far, mobile phone text message reminders may be a potential method in health care systems. Given the widespread use of mobile phone text message reminders among different patient groups, it may have the potential to improve adherence to medication and attendance at clinical appointments globally.

Acknowledgments

This work was in part supported with grants from the Academy of Finland (132581), Turku University Hospital (EVO 13893), Satakunta Hospital District (EVO 12/2010, 81096), Professor Pool and Finnish Cultural Foundation, and the University of Turku.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Descriptive characteristics of the included studies.

[PDF File (Adobe PDF File), 85KB - [jmir_v16i10e222_app1.pdf](#)]

Multimedia Appendix 2

Population, sample size, response rates, description of the intervention, goal of the text messages, dose, and timing of the SMS intervention.

[PDF File (Adobe PDF File), 156KB - [jmir_v16i10e222_app2.pdf](#)]

Multimedia Appendix 3

Assessment methods and outcomes.

[PDF File (Adobe PDF File), 143KB - [jmir_v16i10e222_app3.pdf](#)]

References

1. International Telecommunication Union. ICT Facts and Figures. The World in 2013.: International Telecommunication Union; 2013. URL: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf> [accessed 2014-01-05] [WebCite Cache ID 6MNa1SopR]
2. Atun RF, Sittampalam SR. The role of mobile phones in increasing accessibility and efficiency in healthcare. A Review of the Characteristics and Benefits of SMS in Delivering Healthcare.: The Vodafone Policy Paper Series; 2006. URL: http://www.enlightenmentconomics.com/assets/vodafone_policy_paper_4_march06.pdf [accessed 2014-02-07] [WebCite Cache ID 6ND5OouRy]
3. Zurovac D, Sudoi RK, Akhwale WS, Ndiritu M, Hamer DH, Rowe AK, et al. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: a cluster randomised trial. *Lancet* 2011 Aug 27;378(9793):795-803 [FREE Full text] [doi: [10.1016/S0140-6736\(11\)60783-6](https://doi.org/10.1016/S0140-6736(11)60783-6)] [Medline: [21820166](https://pubmed.ncbi.nlm.nih.gov/21820166/)]
4. Déglise C, Suggs LS, Odermatt P. Short message service (SMS) applications for disease prevention in developing countries. *J Med Internet Res* 2012;14(1):e3 [FREE Full text] [doi: [10.2196/jmir.1823](https://doi.org/10.2196/jmir.1823)] [Medline: [22262730](https://pubmed.ncbi.nlm.nih.gov/22262730/)]

5. Sanghara H, Kravariti E, Jakobsen H, Okocha C. Using short message services in mental health services: assessing feasibility. *Mental Health Review Journal* 2010;15(2):28-33. [doi: [10.5042/mhrj.2010.0369](https://doi.org/10.5042/mhrj.2010.0369)]
6. Finnish Data Protection Ombudsman. Sahkopostin ja tekstiviestien käyttäminen terveydenhuollossa. Tietosuojavaltuutetun kannanotto 1.7.2010. Dnro 1475/41/2009.: Finnish Data Protection Ombudsman; 2010. URL: <http://www.tietosuoja.fi/fi/index/ratkaisut/sahkopostinjatekstiviestienkayttamine.html> [accessed 2014-01-04] [WebCite Cache ID 6MNabXsDI]
7. World Health Organization. National eHealth Strategy Toolkit.: World Health Organization; 2012. URL: http://apps.who.int/iris/bitstream/10665/75211/1/9789241548465_eng.pdf?ua=1 [accessed 2014-02-07] [WebCite Cache ID 6ND4rmmva]
8. European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century.: European Commission; 2012. URL: http://ec.europa.eu/health/ehealth/docs/com_2012_736_en.pdf [accessed 2014-02-07] [WebCite Cache ID 6ND3aPMsr]
9. International mHealth Standards Consortium. Regulatory guidance impacting mHealth Technology.: International mHealth Standards Consortium; 2014. URL: http://www.imhsc.org/regulatory_5.html [accessed 2014-02-07] [WebCite Cache ID 6ND4PG40b]
10. American Telemedicine Association. Telehealth practice recommendations for diabetic retinopathy.: American Telemedicine Association; 2011. URL: <http://www.americantelemed.org/docs/default-source/standards/telehealth-practice-recommendations-for-diabetic-retinopathy.pdf?sfvrsn=10> [accessed 2014-02-07] [WebCite Cache ID 6ND3aPMsr]
11. Kilbourne AM, Neumann MS, Pincus HA, Bauer MS, Stall R. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implement Sci* 2007;2:42 [FREE Full text] [doi: [10.1186/1748-5908-2-42](https://doi.org/10.1186/1748-5908-2-42)] [Medline: [18067681](https://pubmed.ncbi.nlm.nih.gov/18067681/)]
12. Car J, Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R. Mobile phone messaging reminders for attendance at healthcare appointments. *Cochrane Database Syst Rev* 2012;7:CD007458. [doi: [10.1002/14651858.CD007458.pub2](https://doi.org/10.1002/14651858.CD007458.pub2)] [Medline: [22786507](https://pubmed.ncbi.nlm.nih.gov/22786507/)]
13. Vervloet M, Linn AJ, van Weert JC, de Bakker DH, Bouvy ML, van Dijk L. The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: a systematic review of the literature. *J Am Med Inform Assoc* 2012;19(5):696-704 [FREE Full text] [doi: [10.1136/amiajnl-2011-000748](https://doi.org/10.1136/amiajnl-2011-000748)] [Medline: [22534082](https://pubmed.ncbi.nlm.nih.gov/22534082/)]
14. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. *Cochrane Database Syst Rev* 2012;3:CD009756. [doi: [10.1002/14651858.CD009756](https://doi.org/10.1002/14651858.CD009756)] [Medline: [22419345](https://pubmed.ncbi.nlm.nih.gov/22419345/)]
15. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. *Cochrane Database Syst Rev* 2012;12:CD007459. [doi: [10.1002/14651858.CD007459.pub2](https://doi.org/10.1002/14651858.CD007459.pub2)] [Medline: [23235644](https://pubmed.ncbi.nlm.nih.gov/23235644/)]
16. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. *Telemed J E Health* 2009 Apr;15(3):231-240. [doi: [10.1089/tmj.2008.0099](https://doi.org/10.1089/tmj.2008.0099)] [Medline: [19382860](https://pubmed.ncbi.nlm.nih.gov/19382860/)]
17. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for communicating results of medical investigations. *Cochrane Database Syst Rev* 2012;13(6):CD007456. [doi: [10.1002/14651858.CD007456.pub2](https://doi.org/10.1002/14651858.CD007456.pub2)] [Medline: [22696369](https://pubmed.ncbi.nlm.nih.gov/22696369/)]
18. Kauppi K, Välimäki M, Hätönen H, Kuosmanen LM, Warwick-Smith K, Adams CE. Information and communication technology based prompting for treatment compliance for people with serious mental illness. *Cochrane Database Syst Rev* 2014;17(6):CD009960. [doi: [10.1002/14651858.CD009960.pub2](https://doi.org/10.1002/14651858.CD009960.pub2)] [Medline: [24934254](https://pubmed.ncbi.nlm.nih.gov/24934254/)]
19. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009 Jun;26(2):91-108. [doi: [10.1111/j.1471-1842.2009.00848.x](https://doi.org/10.1111/j.1471-1842.2009.00848.x)] [Medline: [19490148](https://pubmed.ncbi.nlm.nih.gov/19490148/)]
20. Nundy S, Razi RR, Dick JJ, Smith B, Mayo A, O'Connor A, et al. A text messaging intervention to improve heart failure self-management after hospital discharge in a largely African-American population: before-after study. *J Med Internet Res* 2013;15(3):e53 [FREE Full text] [doi: [10.2196/jmir.2317](https://doi.org/10.2196/jmir.2317)] [Medline: [23478028](https://pubmed.ncbi.nlm.nih.gov/23478028/)]
21. Furberg RD, Uhrig JD, Bann CM, Lewis MA, Harris JL, Williams P, et al. Technical Implementation of a Multi-Component, Text Message-Based Intervention for Persons Living with HIV. *JMIR Res Protoc* 2012;1(2):e17 [FREE Full text] [doi: [10.2196/resprot.2017](https://doi.org/10.2196/resprot.2017)] [Medline: [23612237](https://pubmed.ncbi.nlm.nih.gov/23612237/)]
22. Shaw RJ, Bosworth HB, Hess JC, Silva SG, Lipkus IM, Davis LL, et al. Development of a Theoretically Driven mHealth Text Messaging Application for Sustaining Recent Weight Loss. *JMIR Mhealth Uhealth* 2013;1(1):e5 [FREE Full text] [doi: [10.2196/mhealth.2343](https://doi.org/10.2196/mhealth.2343)] [Medline: [25100678](https://pubmed.ncbi.nlm.nih.gov/25100678/)]
23. World Economic Forum. The global information technology report.: World Economic Forum; 2013. URL: http://www3.weforum.org/docs/WEF_GITR_Report_2013.pdf [accessed 2014-02-07] [WebCite Cache ID 6ND53iQ0U]
24. CRD. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. York: University of York; 2009. URL: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf [accessed 2014-02-07] [WebCite Cache ID 6ND5DY0TY]
25. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009;6(6):e1000097. [doi: [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)]

26. 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](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
27. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res* 2012;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](https://doi.org/10.2196/jmir.2104)] [Medline: [23151820](https://pubmed.ncbi.nlm.nih.gov/23151820/)]
28. Rush KL, Adamack M, Gordon J, Lilly M, Janke R. Best practices of formal new graduate nurse transition programs: an integrative review. *Int J Nurs Stud* 2013 Mar;50(3):345-356. [doi: [10.1016/j.ijnurstu.2012.06.009](https://doi.org/10.1016/j.ijnurstu.2012.06.009)] [Medline: [22795800](https://pubmed.ncbi.nlm.nih.gov/22795800/)]
29. Pluye P, Robert E, Cargo M, Bartlett G, O’Cathain A, Griffiths F, et al. Proposal: A mixed methods appraisal tool for systematic mixed studies reviews. 2011. URL: <http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/attach/84371689/MMAT%202011%20criteria%20and%20tutorial%202011-06-29updated2014.08.21.pdf> [accessed 2014-01-04] [WebCite Cache ID 6MNBseB0r]
30. Pijnenborg GHM, Withaar FK, Evans JJ, van den Bosch RJ, Brouwer WH. SMS text messages as a prosthetic aid in the cognitive rehabilitation of schizophrenia. *Rehabilitation Psychology* 2007;52(2):236-240. [doi: [10.1037/0090-5550.52.2.236](https://doi.org/10.1037/0090-5550.52.2.236)]
31. Pena-Robichaux V, Kvedar JC, Watson AJ. Text messages as a reminder aid and educational tool in adults and adolescents with atopic dermatitis: a pilot study. *Dermatol Res Pract* 2010;2010:- [FREE Full text] [doi: [10.1155/2010/894258](https://doi.org/10.1155/2010/894258)] [Medline: [20885940](https://pubmed.ncbi.nlm.nih.gov/20885940/)]
32. Arora S, Peters AL, Agy C, Menchine M. A mobile health intervention for inner city patients with poorly controlled diabetes: proof-of-concept of the TExT-MED program. *Diabetes Technol Ther* 2012 Jun;14(6):492-496. [doi: [10.1089/dia.2011.0252](https://doi.org/10.1089/dia.2011.0252)] [Medline: [22524591](https://pubmed.ncbi.nlm.nih.gov/22524591/)]
33. Lua PL, Neni WS. Feasibility and acceptability of mobile epilepsy educational system (MEES) for people with epilepsy in Malaysia. *Telemed J E Health* 2012 Dec;18(10):777-784 [FREE Full text] [doi: [10.1089/tmj.2012.0047](https://doi.org/10.1089/tmj.2012.0047)] [Medline: [23078181](https://pubmed.ncbi.nlm.nih.gov/23078181/)]
34. Lund S, Hemed M, Nielsen BB, Said A, Said K, Makungu MH, et al. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: a cluster-randomised controlled trial. *BJOG* 2012 Sep;119(10):1256-1264. [doi: [10.1111/j.1471-0528.2012.03413.x](https://doi.org/10.1111/j.1471-0528.2012.03413.x)] [Medline: [22805598](https://pubmed.ncbi.nlm.nih.gov/22805598/)]
35. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. *Diabet Med* 2006 Dec;23(12):1332-1338. [doi: [10.1111/j.1464-5491.2006.01989.x](https://doi.org/10.1111/j.1464-5491.2006.01989.x)] [Medline: [17116184](https://pubmed.ncbi.nlm.nih.gov/17116184/)]
36. Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. *Diabetes Technol Ther* 2009 Feb;11(2):99-106. [doi: [10.1089/dia.2008.0022](https://doi.org/10.1089/dia.2008.0022)] [Medline: [19848576](https://pubmed.ncbi.nlm.nih.gov/19848576/)]
37. Balato N, Megna M, Di Costanzo L, Balato A, Ayala F. Educational and motivational support service: a pilot study for mobile-phone-based interventions in patients with psoriasis. *Br J Dermatol* 2013 Jan;168(1):201-205. [doi: [10.1111/j.1365-2133.2012.11205.x](https://doi.org/10.1111/j.1365-2133.2012.11205.x)] [Medline: [23240729](https://pubmed.ncbi.nlm.nih.gov/23240729/)]
38. Agyapong VI, Milnes J, McLoughlin DM, Farren CK. Perception of patients with alcohol use disorder and comorbid depression about the usefulness of supportive text messages. *Technol Health Care* 2013;21(1):31-39. [doi: [10.3233/THC-120707](https://doi.org/10.3233/THC-120707)] [Medline: [23358057](https://pubmed.ncbi.nlm.nih.gov/23358057/)]
39. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. *J Med Internet Res* 2004 Dec 2;6(4):e42 [FREE Full text] [doi: [10.2196/jmir.6.4.e42](https://doi.org/10.2196/jmir.6.4.e42)] [Medline: [15631966](https://pubmed.ncbi.nlm.nih.gov/15631966/)]
40. Rodrigues R, Shet A, Antony J, Sidney K, Arumugam K, Krishnamurthy S, et al. Supporting adherence to antiretroviral therapy with mobile phone reminders: results from a cohort in South India. *PLoS One* 2012;7(8):e40723 [FREE Full text] [doi: [10.1371/journal.pone.0040723](https://doi.org/10.1371/journal.pone.0040723)] [Medline: [22952574](https://pubmed.ncbi.nlm.nih.gov/22952574/)]
41. Sidney K, Antony J, Rodrigues R, Arumugam K, Krishnamurthy S, D'souza G, et al. Supporting patient adherence to antiretrovirals using mobile phone reminders: patient responses from South India. *AIDS Care* 2012;24(5):612-617. [doi: [10.1080/09540121.2011.630357](https://doi.org/10.1080/09540121.2011.630357)] [Medline: [22150088](https://pubmed.ncbi.nlm.nih.gov/22150088/)]
42. Hou MY, Hurwitz S, Kavanagh E, Fortin J, Goldberg AB. Using daily text-message reminders to improve adherence with oral contraceptives: a randomized controlled trial. *Obstet Gynecol* 2010 Sep;116(3):633-640. [doi: [10.1097/AOG.0b013e3181eb6b0f](https://doi.org/10.1097/AOG.0b013e3181eb6b0f)] [Medline: [20733446](https://pubmed.ncbi.nlm.nih.gov/20733446/)]
43. Strandbygaard U, Thomsen SF, Backer V. A daily SMS reminder increases adherence to asthma treatment: a three-month follow-up study. *Respir Med* 2010 Feb;104(2):166-171 [FREE Full text] [doi: [10.1016/j.rmed.2009.10.003](https://doi.org/10.1016/j.rmed.2009.10.003)] [Medline: [19854632](https://pubmed.ncbi.nlm.nih.gov/19854632/)]
44. Dick JJ, Nundy S, Solomon MC, Bishop KN, Chin MH, Peek ME. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. *J Diabetes Sci Technol* 2011 Sep;5(5):1246-1254 [FREE Full text] [Medline: [22027326](https://pubmed.ncbi.nlm.nih.gov/22027326/)]
45. Hardy H, Kumar V, Doros G, Farmer E, Drainoni ML, Rybin D, et al. Randomized controlled trial of a personalized cellular phone reminder system to enhance adherence to antiretroviral therapy. *AIDS Patient Care STDS* 2011 Mar;25(3):153-161 [FREE Full text] [doi: [10.1089/apc.2010.0006](https://doi.org/10.1089/apc.2010.0006)] [Medline: [21323532](https://pubmed.ncbi.nlm.nih.gov/21323532/)]

46. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, de Walque D, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS* 2011 Mar 27;25(6):825-834 [FREE Full text] [doi: [10.1097/QAD.0b013e32834380c1](https://doi.org/10.1097/QAD.0b013e32834380c1)] [Medline: [21252632](https://pubmed.ncbi.nlm.nih.gov/21252632/)]
47. Boker A, Feetham HJ, Armstrong A, Purcell P, Jacobe H. Do automated text messages increase adherence to acne therapy? Results of a randomized, controlled trial. *J Am Acad Dermatol* 2012 Dec;67(6):1136-1142. [doi: [10.1016/j.jaad.2012.02.031](https://doi.org/10.1016/j.jaad.2012.02.031)] [Medline: [22521201](https://pubmed.ncbi.nlm.nih.gov/22521201/)]
48. Dowshen N, Kuhns LM, Johnson A, Holoyda BJ, Garofalo R. Improving adherence to antiretroviral therapy for youth living with HIV/AIDS: a pilot study using personalized, interactive, daily text message reminders. *J Med Internet Res* 2012;14(2):e51 [FREE Full text] [doi: [10.2196/jmir.2015](https://doi.org/10.2196/jmir.2015)] [Medline: [22481246](https://pubmed.ncbi.nlm.nih.gov/22481246/)]
49. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): a pilot trial of an interactive text-messaging intervention for medication adherence, socialization, and auditory hallucinations. *Schizophr Bull* 2012 May;38(3):414-425 [FREE Full text] [doi: [10.1093/schbul/sbr155](https://doi.org/10.1093/schbul/sbr155)] [Medline: [22080492](https://pubmed.ncbi.nlm.nih.gov/22080492/)]
50. Lv Y, Zhao H, Liang Z, Dong H, Liu L, Zhang D, et al. A mobile phone short message service improves perceived control of asthma: a randomized controlled trial. *Telemed J E Health* 2012;18(6):420-426. [doi: [10.1089/tmj.2011.0218](https://doi.org/10.1089/tmj.2011.0218)] [Medline: [22667695](https://pubmed.ncbi.nlm.nih.gov/22667695/)]
51. Montes JM, Medina E, Gomez-Beneyto M, Maurino J. A short message service (SMS)-based strategy for enhancing adherence to antipsychotic medication in schizophrenia. *Psychiatry Res* 2012 Dec 30;200(2-3):89-95. [doi: [10.1016/j.psychres.2012.07.034](https://doi.org/10.1016/j.psychres.2012.07.034)] [Medline: [22901437](https://pubmed.ncbi.nlm.nih.gov/22901437/)]
52. Ting TV, Kudalkar D, Nelson S, Cortina S, Pendl J, Budhani S, et al. Usefulness of cellular text messaging for improving adherence among adolescents and young adults with systemic lupus erythematosus. *J Rheumatol* 2012 Jan;39(1):174-179. [doi: [10.3899/jrheum.110771](https://doi.org/10.3899/jrheum.110771)] [Medline: [22089460](https://pubmed.ncbi.nlm.nih.gov/22089460/)]
53. Lewis MA, Uhrig JD, Bann CM, Harris JL, Furberg RD, Coomes C, et al. Tailored text messaging intervention for HIV adherence: a proof-of-concept study. *Health Psychol* 2013 Mar;32(3):248-253. [doi: [10.1037/a0028109](https://doi.org/10.1037/a0028109)] [Medline: [22545972](https://pubmed.ncbi.nlm.nih.gov/22545972/)]
54. Nundy S, Dick JJ, Solomon MC, Peek ME. Developing a behavioral model for mobile phone-based diabetes interventions. *Patient Educ Couns* 2013 Jan;90(1):125-132 [FREE Full text] [doi: [10.1016/j.pec.2012.09.008](https://doi.org/10.1016/j.pec.2012.09.008)] [Medline: [23063349](https://pubmed.ncbi.nlm.nih.gov/23063349/)]
55. Mao Y, Zhang Y, Zhai S. Mobile phone text messaging for pharmaceutical care in a hospital in China. *J Telemed Telecare* 2008;14(8):410-414. [doi: [10.1258/jtt.2008.080406](https://doi.org/10.1258/jtt.2008.080406)] [Medline: [19047450](https://pubmed.ncbi.nlm.nih.gov/19047450/)]
56. Brannan SO, Dewar C, Taggerty L, Clark S. The effect of short messaging service text on non-attendance in a general ophthalmology clinic. *Scott Med J* 2011 Aug;56(3):148-150. [doi: [10.1258/smj.2011.011112](https://doi.org/10.1258/smj.2011.011112)] [Medline: [21873719](https://pubmed.ncbi.nlm.nih.gov/21873719/)]
57. Vilella A, Bayas JM, Diaz MT, Guinovart C, Diez C, Simó D, et al. The role of mobile phones in improving vaccination rates in travelers. *Prev Med* 2004 Apr;38(4):503-509. [doi: [10.1016/j.ypmed.2003.12.005](https://doi.org/10.1016/j.ypmed.2003.12.005)] [Medline: [15020186](https://pubmed.ncbi.nlm.nih.gov/15020186/)]
58. Bos A, Hoogstraten J, Prah-Andersen B. Failed appointments in an orthodontic clinic. *Am J Orthod Dentofacial Orthop* 2005 Mar;127(3):355-357. [doi: [10.1016/j.ajodo.2004.11.014](https://doi.org/10.1016/j.ajodo.2004.11.014)] [Medline: [15775951](https://pubmed.ncbi.nlm.nih.gov/15775951/)]
59. Downer SR, Meara JG, Da Costa AC. Use of SMS text messaging to improve outpatient attendance. *Med J Aust* 2005 Oct 3;183(7):366-368. [Medline: [16201955](https://pubmed.ncbi.nlm.nih.gov/16201955/)]
60. Leong KC, Chen WS, Leong KW, Mastura I, Mimi O, Sheikh MA, et al. The use of text messaging to improve attendance in primary care: a randomized controlled trial. *Fam Pract* 2006 Dec;23(6):699-705 [FREE Full text] [doi: [10.1093/fampra/cml044](https://doi.org/10.1093/fampra/cml044)] [Medline: [16916871](https://pubmed.ncbi.nlm.nih.gov/16916871/)]
61. Chen ZW, Fang LZ, Chen LY, Dai HL. Comparison of an SMS text messaging and phone reminder to improve attendance at a health promotion center: a randomized controlled trial. *J Zhejiang Univ Sci B* 2008 Jan;9(1):34-38 [FREE Full text] [doi: [10.1631/jzus.B071464](https://doi.org/10.1631/jzus.B071464)] [Medline: [18196610](https://pubmed.ncbi.nlm.nih.gov/18196610/)]
62. Fairhurst K, Sheikh A. Texting appointment reminders to repeated non-attenders in primary care: randomised controlled study. *Qual Saf Health Care* 2008 Oct;17(5):373-376. [doi: [10.1136/qshc.2006.020768](https://doi.org/10.1136/qshc.2006.020768)] [Medline: [18842978](https://pubmed.ncbi.nlm.nih.gov/18842978/)]
63. Koshy E, Car J, Majeed A. Effectiveness of mobile-phone short message service (SMS) reminders for ophthalmology outpatient appointments: observational study. *BMC Ophthalmol* 2008;8:9 [FREE Full text] [doi: [10.1186/1471-2415-8-9](https://doi.org/10.1186/1471-2415-8-9)] [Medline: [18513438](https://pubmed.ncbi.nlm.nih.gov/18513438/)]
64. Foley J, O'Neill M. Use of mobile telephone short message service (SMS) as a reminder: the effect on patient attendance. *Eur Arch Paediatr Dent* 2009 Jan;10(1):15-18. [Medline: [19254521](https://pubmed.ncbi.nlm.nih.gov/19254521/)]
65. Liew SM, Tong SF, Lee VK, Ng CJ, Leong KC, Teng CL. Text messaging reminders to reduce non-attendance in chronic disease follow-up: a clinical trial. *Br J Gen Pract* 2009 Dec;59(569):916-920 [FREE Full text] [doi: [10.3399/bjgp09X472250](https://doi.org/10.3399/bjgp09X472250)] [Medline: [19712544](https://pubmed.ncbi.nlm.nih.gov/19712544/)]
66. Ludlow H, Hurley J, Dolwani S. Using email and text messaging to improve patient compliance with blood monitoring. *Nurs Times* 2009;105(28):26-28. [Medline: [19715236](https://pubmed.ncbi.nlm.nih.gov/19715236/)]
67. da Costa TM, Salomão P, Martha AS, Pisa IT, Sigulem D. The impact of short message service text messages sent as appointment reminders to patients' cell phones at outpatient clinics in São Paulo, Brazil. *Int J Med Inform* 2010 Jan;79(1):65-70. [doi: [10.1016/j.ijmedinf.2009.09.001](https://doi.org/10.1016/j.ijmedinf.2009.09.001)] [Medline: [19783204](https://pubmed.ncbi.nlm.nih.gov/19783204/)]
68. Milne RG. Reducing non-attendance at specialist clinics: an evaluation of the effectiveness and cost of patient-focused booking and SMS reminders at a Scottish health board. *Int J Consum Stud* 2010;34(5):570-580. [doi: [10.1111/j.1470-6431.2010.00903.x](https://doi.org/10.1111/j.1470-6431.2010.00903.x)]

69. Bourne C, Knight V, Guy R, Wand H, Lu H, McNulty A. Short message service reminder intervention doubles sexually transmitted infection/HIV re-testing rates among men who have sex with men. *Sex Transm Infect* 2011 Apr;87(3):229-231. [doi: [10.1136/sti.2010.048397](https://doi.org/10.1136/sti.2010.048397)] [Medline: [21296796](https://pubmed.ncbi.nlm.nih.gov/21296796/)]
70. Branson CE, Clemmey P, Mukherjee P. Text message reminders to improve outpatient therapy attendance among adolescents: a pilot study. *Psychol Serv* 2013 Aug;10(3):298-303. [doi: [10.1037/a0026693](https://doi.org/10.1037/a0026693)] [Medline: [23937089](https://pubmed.ncbi.nlm.nih.gov/23937089/)]
71. Perry JG. A preliminary investigation into the effect of the use of the Short Message Service (SMS) on patient attendance at an NHS Dental Access Centre in Scotland. *Prim Dent Care* 2011 Oct;18(4):145-149. [doi: [10.1308/135576111797512810](https://doi.org/10.1308/135576111797512810)] [Medline: [21968040](https://pubmed.ncbi.nlm.nih.gov/21968040/)]
72. Altuwaijri MM, Sughayr AM, Hassan MA, Alazwari FM. The effect of integrating short messaging services` reminders with electronic medical records on non-attendance rates. *Saudi Med J* 2012 Feb;33(2):193-196. [Medline: [22327762](https://pubmed.ncbi.nlm.nih.gov/22327762/)]
73. Fischer HH, Moore SL, Ginosar D, Davidson AJ, Rice-Peterson CM, Durfee MJ, et al. Care by cell phone: text messaging for chronic disease management. *Am J Manag Care* 2012 Feb;18(2):e42-e47 [FREE Full text] [Medline: [22435883](https://pubmed.ncbi.nlm.nih.gov/22435883/)]
74. Prasad S, Anand R. Use of mobile telephone short message service as a reminder: the effect on patient attendance. *Int Dent J* 2012 Feb;62(1):21-26. [doi: [10.1111/j.1875-595X.2011.00081.x](https://doi.org/10.1111/j.1875-595X.2011.00081.x)] [Medline: [22251033](https://pubmed.ncbi.nlm.nih.gov/22251033/)]
75. Sims H, Sanghara H, Hayes D, Wandiembe S, Finch M, Jakobsen H, et al. Text message reminders of appointments: a pilot intervention at four community mental health clinics in London. *Psychiatr Serv* 2012 Feb 1;63(2):161-168. [doi: [10.1176/appi.ps.201100211](https://doi.org/10.1176/appi.ps.201100211)] [Medline: [22302334](https://pubmed.ncbi.nlm.nih.gov/22302334/)]
76. Taylor NF, Bottrell J, Lawler K, Benjamin D. Mobile telephone short message service reminders can reduce nonattendance in physical therapy outpatient clinics: a randomized controlled trial. *Arch Phys Med Rehabil* 2012 Jan;93(1):21-26. [doi: [10.1016/j.apmr.2011.08.007](https://doi.org/10.1016/j.apmr.2011.08.007)] [Medline: [22000821](https://pubmed.ncbi.nlm.nih.gov/22000821/)]
77. Downing SG, Cashman C, McNamee H, Penney D, Russell DB, Hellard ME. Increasing chlamydia test of re-infection rates using SMS reminders and incentives. *Sex Transm Infect* 2013 Feb;89(1):16-19. [doi: [10.1136/sextrans-2011-050454](https://doi.org/10.1136/sextrans-2011-050454)] [Medline: [22728911](https://pubmed.ncbi.nlm.nih.gov/22728911/)]
78. Guy R, Wand H, Knight V, Kenigsberg A, Read P, McNulty AM. SMS reminders improve re-screening in women and heterosexual men with chlamydia infection at Sydney Sexual Health Centre: a before-and-after study. *Sex Transm Infect* 2013 Feb;89(1):11-15. [doi: [10.1136/sextrans-2011-050370](https://doi.org/10.1136/sextrans-2011-050370)] [Medline: [22517890](https://pubmed.ncbi.nlm.nih.gov/22517890/)]
79. Kollmann A, Riedl M, Kastner P, Schreier G, Ludvik B. Feasibility of a mobile phone-based data service for functional insulin treatment of type 1 diabetes mellitus patients. *J Med Internet Res* 2007;9(5):e36 [FREE Full text] [doi: [10.2196/jmir.9.5.e36](https://doi.org/10.2196/jmir.9.5.e36)] [Medline: [18166525](https://pubmed.ncbi.nlm.nih.gov/18166525/)]
80. da Costa TM, Barbosa BJ, Gomes e Costa DA, Sigulem D, de Fátima Marin H, Filho AC, et al. Results of a randomized controlled trial to assess the effects of a mobile SMS-based intervention on treatment adherence in HIV/AIDS-infected Brazilian women and impressions and satisfaction with respect to incoming messages. *Int J Med Inform* 2012 Apr;81(4):257-269 [FREE Full text] [doi: [10.1016/j.ijmedinf.2011.10.002](https://doi.org/10.1016/j.ijmedinf.2011.10.002)] [Medline: [22296762](https://pubmed.ncbi.nlm.nih.gov/22296762/)]
81. Vervloet M, van Dijk L, Santen-Reestman J, van Vlijmen B, van Wingerden P, Bouvy ML, et al. SMS reminders improve adherence to oral medication in type 2 diabetes patients who are real time electronically monitored. *Int J Med Inform* 2012 Sep;81(9):594-604. [doi: [10.1016/j.ijmedinf.2012.05.005](https://doi.org/10.1016/j.ijmedinf.2012.05.005)] [Medline: [22652012](https://pubmed.ncbi.nlm.nih.gov/22652012/)]
82. Foreman KF, Stockl KM, Le LB, Fisk E, Shah SM, Lew HC, et al. Impact of a text messaging pilot program on patient medication adherence. *Clin Ther* 2012 May;34(5):1084-1091. [doi: [10.1016/j.clinthera.2012.04.007](https://doi.org/10.1016/j.clinthera.2012.04.007)] [Medline: [22554973](https://pubmed.ncbi.nlm.nih.gov/22554973/)]
83. Pijnenborg GH, Withaar FK, Brouwer WH, Timmerman ME, van den Bosch RJ, Evans JJ. The efficacy of SMS text messages to compensate for the effects of cognitive impairments in schizophrenia. *Br J Clin Psychol* 2010 Jun;49(Pt 2):259-274. [doi: [10.1348/014466509X467828](https://doi.org/10.1348/014466509X467828)] [Medline: [19735607](https://pubmed.ncbi.nlm.nih.gov/19735607/)]
84. Holtz B, Whitten P. Managing asthma with mobile phones: a feasibility study. *Telemed J E Health* 2009 Nov;15(9):907-909. [doi: [10.1089/tmj.2009.0048](https://doi.org/10.1089/tmj.2009.0048)] [Medline: [19919198](https://pubmed.ncbi.nlm.nih.gov/19919198/)]
85. Britto MT, Munafò JK, Schoettker PJ, Vockell AL, Wimberg JA, Yi MS. Pilot and feasibility test of adolescent-controlled text messaging reminders. *Clin Pediatr (Phila)* 2012 Feb;51(2):114-121. [doi: [10.1177/0009922811412950](https://doi.org/10.1177/0009922811412950)] [Medline: [22019792](https://pubmed.ncbi.nlm.nih.gov/22019792/)]
86. Greaney ML, Puleo E, Sprunck-Harrild K, Bennett GG, Cunningham MA, Gillman MW, et al. Electronic reminders for cancer prevention: factors associated with preference for automated voice reminders or text messages. *Prev Med* 2012 Aug;55(2):151-154 [FREE Full text] [doi: [10.1016/j.yjmed.2012.05.014](https://doi.org/10.1016/j.yjmed.2012.05.014)] [Medline: [22659227](https://pubmed.ncbi.nlm.nih.gov/22659227/)]
87. Guy R, Hocking J, Wand H, Stott S, Ali H, Kaldor J. How effective are short message service reminders at increasing clinic attendance? A meta-analysis and systematic review. *Health Serv Res* 2012 Apr;47(2):614-632 [FREE Full text] [doi: [10.1111/j.1475-6773.2011.01342.x](https://doi.org/10.1111/j.1475-6773.2011.01342.x)] [Medline: [22091980](https://pubmed.ncbi.nlm.nih.gov/22091980/)]
88. Whittaker R. Issues in mHealth: findings from key informant interviews. *J Med Internet Res* 2012;14(5):e129 [FREE Full text] [doi: [10.2196/jmir.1989](https://doi.org/10.2196/jmir.1989)] [Medline: [23032424](https://pubmed.ncbi.nlm.nih.gov/23032424/)]
89. Ahlers-Schmidt CR, Chesser AK, Nguyen T, Brannon J, Hart TA, Williams KS, et al. Feasibility of a randomized controlled trial to evaluate Text Reminders for Immunization Compliance in Kids (TRICKS). *Vaccine* 2012 Aug 3;30(36):5305-5309. [doi: [10.1016/j.vaccine.2012.06.058](https://doi.org/10.1016/j.vaccine.2012.06.058)] [Medline: [22750044](https://pubmed.ncbi.nlm.nih.gov/22750044/)]

90. Wakadha H, Chandir S, Were EV, Rubin A, Obor D, Levine OS, et al. The feasibility of using mobile-phone based SMS reminders and conditional cash transfers to improve timely immunization in rural Kenya. *Vaccine* 2013 Jan 30;31(6):987-993. [doi: [10.1016/j.vaccine.2012.11.093](https://doi.org/10.1016/j.vaccine.2012.11.093)] [Medline: [23246258](https://pubmed.ncbi.nlm.nih.gov/23246258/)]

Abbreviations

ICT: information and communication technology

MMAT: Mixed Method Appraisal Tool

RCT: randomized controlled trial

SMS: short message service

Edited by G Eysenbach; submitted 02.04.14; peer-reviewed by H Sanghara, B Balzer; comments to author 17.07.14; revised version received 30.07.14; accepted 25.08.14; published 17.10.14.

Please cite as:

Kannisto KA, Koivunen MH, Välimäki MA

Use of Mobile Phone Text Message Reminders in Health Care Services: A Narrative Literature Review

J Med Internet Res 2014;16(10):e222

URL: <http://www.jmir.org/2014/10/e222/>

doi: [10.2196/jmir.3442](https://doi.org/10.2196/jmir.3442)

PMID: [25326646](https://pubmed.ncbi.nlm.nih.gov/25326646/)

©Kati Anneli Kannisto, Marita Hannele Koivunen, Maritta Anneli Välimäki. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 17.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Beyond Traditional Advertisements: Leveraging Facebook's Social Structures for Research Recruitment

Rupa S Valdez¹, PhD; Thomas M Guterbock^{1,2,3}, PhD; Morgan J Thompson¹, BA; Jeremiah D Reilly¹; Hannah K Menefee¹, BA; Maria S Bennici^{1,4}, BA; Ishan C Williams⁵, PhD; Deborah L Rexrode², PhD

¹Department of Public Health Sciences, University of Virginia, Charlottesville, VA, United States

²Center for Survey Research, University of Virginia, Charlottesville, VA, United States

³Department of Sociology, University of Virginia, Charlottesville, VA, United States

⁴Paris School of International Affairs, Institut d'Etudes Politiques de Paris, Paris, France

⁵School of Nursing, University of Virginia, Charlottesville, VA, United States

Corresponding Author:

Rupa S Valdez, PhD

Department of Public Health Sciences

University of Virginia

PO Box 800717

Charlottesville, VA, 22908

United States

Phone: 1 434 982 2510

Fax: 1 434 243 3418

Email: rupavaldez@virginia.edu

Abstract

Background: Obtaining access to a demographically and geographically diverse sample for health-related research can be costly and time consuming. Previous studies have reported mixed results regarding the potential of using social media-based advertisements to overcome these challenges.

Objective: Our aim was to develop and assess the feasibility, benefits, and challenges of recruiting for research studies related to consumer health information technology (IT) by leveraging the social structures embedded in the social networking platform, Facebook.

Methods: Two recruitment strategies that involved direct communication with existing Facebook groups and pages were developed and implemented in two distinct populations. The first recruitment strategy involved posting a survey link directly to consenting groups and pages and was used to recruit Filipino-Americans to a study assessing the perceptions, use of, and preferences for consumer health IT. This study took place between August and December 2013. The second recruitment strategy targeted individuals with type 2 diabetes and involved creating a study-related Facebook group and asking administrators of other groups and pages to publicize our group to their members. Group members were then directly invited to participate in an online pre-study survey. This portion of a larger study to understand existing health management practices as a foundation for consumer health IT design took place between May and June 2014. In executing both recruitment strategies, efforts were made to establish trust and transparency. Recruitment rate, cost, content of interaction, and characteristics of the sample obtained were used to assess the recruitment methods.

Results: The two recruitment methods yielded 87 and 79 complete responses, respectively. The first recruitment method yielded a rate of study completion proportionate to that of the rate of posts made, whereas recruitment successes of the second recruitment method seemed to follow directly from the actions of a subset of administrators. Excluding personnel time, the first recruitment method resulted in no direct costs, and the second recruitment method resulted in a total direct cost of US \$118.17. Messages, posts, and comments received using both recruitment strategies reflected ten themes, including appreciation, assistance, clarification, concerns, encouragement, health information, interest, promotion, solicitations, and support. Both recruitment methods produced mixed results regarding sample representativeness with respect to characteristics such as gender, race, and ethnicity.

Conclusions: The results of the study demonstrate that leveraging the social structures of Facebook for health-related research was feasible for obtaining small samples appropriate for qualitative research but not for obtaining large samples needed for quantitative research. The content of interactions with members of the target population prompted ethical deliberations concerning

suitable target communities and appropriate boundaries between researchers and participants. Widespread replication of this method would benefit from a broad discussion among researchers, social media users, social media companies, and experts in research ethics to address appropriate protocols for such interactions.

(*J Med Internet Res* 2014;16(10):e243) doi:[10.2196/jmir.3786](https://doi.org/10.2196/jmir.3786)

KEYWORDS

participant recruitment; Facebook; social media; consumer health IT; ethnicity; advertising

Introduction

As consumer health information technology (IT) becomes increasingly integral to the health care delivery system, it is imperative that such technology aligns with patients' needs [1-4]. Gaining in-depth understanding of these requirements necessitates directly assessing patients' existing health management practices [5,6] and experiences with prototypes of and currently available consumer health IT solutions [2,7]. Yet, obtaining in-person access to patient populations can be challenging, and researchers' and designers' recruitment efforts are often limited to geographically available populations due to time and cost. Consequently, research questions must fit available populations, and recruitment efforts may fail to attain populations with targeted characteristics related to race, ethnicity, and socioeconomic status. Due to their size and diversity of users, online social media platforms offer the potential to engage a wider range of patient populations in consumer health IT research.

This paper's purpose is to present the development and assessment of recruitment strategies leveraging one social media platform, Facebook, for consumer health IT research and design. Facebook enables users to create their own profile pages and connect with others [8,9]. It also enables multiple modes of communication between individuals. Public modes include updating one's status on one's own profile page, posting content on another's profile page, and commenting on others' statuses, posts, or photos. Private modes include messaging (similar to emailing), chatting (similar to instant messaging), and video chatting (similar to video conferencing). In addition, users may create, join, and interact with groups focusing on specific topics and like, post to, and follow pages (public profiles for businesses and organizations). Each user's home page features a news feed, a tailored selection of updates about the user's friends' Facebook activity [10]. Users may modify privacy settings, including specifying with whom information may be shared and from whom contact requests may be received.

Facebook is a promising recruitment instrument for three reasons. First, it remains the most visited social networking site within the United States via computer, mobile phone apps, and mobile phone browsers, with a significantly larger user base and per person usage time than the next largest social networking site across each platform [11]. Facebook reported a total of 1.28 billion active monthly users at the end of the first quarter of 2014 [12], with 71% of online adults in the United States using this service [13]. Moreover, 81.2% of daily active users reside outside the United States and Canada [14]. Consequently, Facebook may facilitate access to an extensive

group of potential participants both within the United States and globally.

Second, within the United States, Facebook boasts a diversity of users [13]. No statistically significant differences exist between the proportions of online adult non-Hispanic whites, non-Hispanic blacks, and Hispanics using Facebook. Similarly, no statistically significant differences exist between the proportions residing in urban, rural, and suburban settings. Although statistically significant differences do exist across socioeconomic status and education, the largest spread between categories does not exceed 8%. The most meaningful differences in usage exist across gender and age; online women and younger adults are more likely to use Facebook. However, across both genders and all age groups except for those over 65, a majority of individuals engage with the platform. While only 45% of those over 65 currently use Facebook, this demographic is one of the fastest growing user groups [15]. Thus, directing recruiting through Facebook has the potential to engage individuals from all demographic categories.

Finally, Facebook is becoming a space for health-related activity and exchange. In 2011, Pew Research Center surveys indicated that 11% of Facebook users had posted comments, queries, or information about health or medical matters, and 9% had started or joined a health-related group on a social networking site [16]. In the same year, one study found over 600 breast cancer groups on Facebook, with a collective membership of over one million [17]. Health-related Facebook groups are used for sharing personal clinical information, requesting disease-specific guidance, receiving emotional support, fundraising, and generating awareness about a condition [17,18]. Additionally, health care organizations, including governmental centers and agencies, health care institutions, pharmaceutical companies, and nonprofits use Facebook to disseminate health advice and promotion messages [19]. Users also communicate about health through their individual profiles [20-23]. Facebook has naturally extended to encompass health-related topics; therefore, using it as a recruitment platform for health IT-related research activities is consistent with its existing scope.

Nonetheless, recruiting research participants via Facebook involves significant and novel obstacles. Encouraging people to engage in research requires more than an invitation to participate. Dillman et al have argued that participation is generated through a successfully negotiated process of social exchange, using techniques they call the "tailored design method" [24]. They note that "Tailored design is the development of survey procedures that work together to form the survey request and motivate various types of people to respond to the survey by establishing trust and increasing the perceived benefits of completing the survey while decreasing

the expected costs of participation” p. 38 [24]. Creating trust and transparency is key to recruitment; however, this can be challenging in any computer-mediated communication, including group interaction. Rules and restrictions enforced by Facebook, as well as informal, emergent norms in online communities, restrict the forms and channels of communication available to researchers. As columnist David Brooks recently suggested, social media may be changing implicit assumptions about which people and which institutions are worthy of trust [25].

Within the health sciences, recruitment efforts for research via Facebook have predominantly used paid advertisements [26-32]. Clients of Facebook’s advertisement services are able to target individuals based on location, interests/hobbies, and information obtained from users’ profiles [33]. This has met with mixed results. Some studies have concluded that recruitment through Facebook advertisements is effective in terms of cost, participant yield, and ability to engage specific demographic groups such as low-income individuals and individuals with specific health conditions [27-30]. However, others have determined that this method yields few participants, or yields them at substantial cost [26,31,32], in some cases resulting in no meaningful participation [31]. This range of outcomes may be partly explained by the target populations, the subject matter of the research, and the actions requested by the researchers (eg, online survey, participation in a clinical trial). All of these variables can affect respondents’ perceptions of trust, legitimacy, benefit, and cost and determine their rate of response [24].

In some cases, therefore, advertisements are effective for recruiting via Facebook. However, the advertisement-based approach only partially capitalizes on Facebook as a platform, as it relies on broadcast strategies. Methods leveraging the social structures embedded within Facebook may provide additional means of recruiting. A few studies in the health sciences have initiated exploration of such methods. Zaid et al posted a link to an online survey to one Facebook group for women with neuroendocrine carcinoma of the cervix [34]. They received 57 survey responses within the 30-day period during which the study was open. Others succeeded by directly communicating from a study or personal Facebook page [32,35]. One study mentioned posting directly to Facebook groups and pages related to the health condition of interest but provided no details about this method or its success [23]. No health science studies we know of have used recruitment strategies that involve directly contacting administrators of multiple groups and pages or creating a study group as a sampling frame.

This paper reports on our experiences implementing two recruitment strategies to leverage the tailored design method by directly communicating with administrators and members of Facebook groups and pages. We specifically address the feasibility, benefits, and challenges of implementing such strategies within the context of consumer health IT research. Additionally, we report on our experiences targeting individuals (1) from a specific ethnic group (ie, Filipino-Americans) and (2) with a specific diagnosis (ie, type 2 diabetes) from multiple demographic groups. To our knowledge, this is the first assessment of such recruitment strategies capitalizing on the

complex social structures of an online social network for health-related research.

Methods

Study Details

Study 1: Consumer Health Information Technology in a Filipino Community

Study 1 was designed to assess the feasibility, benefits, and challenges of using Facebook to recruit members of a specific ethnic group, Filipino-Americans, for a survey assessing the perceptions, use of, and preferences for consumer health IT. Recruitment occurred between August and December 2013. This study focused on Filipino-Americans because, despite Asian Americans’ extensive online presence [36], the needs and preferences for consumer health IT of specific communities within this population remain largely uncharacterized. Participants completed an online survey administered through SurveyMonkey [37], which contained 33 closed and 4 open-ended questions on topics including (1) experiences with two forms of consumer health IT (ie, personal health records and mobile health applications) and general social networking sites for health management, (2) preferences for the design of consumer health IT targeting the Filipino-American community, and (3) demographics. Eligible individuals were 18 years or older, identified as Filipino, and lived in the United States. Prior to launch, the online survey was piloted and revised based on feedback from 8 individuals identifying as Filipino-American.

Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

Study 2 was designed to assess the feasibility, benefits, and challenges of using Facebook to recruit individuals with type 2 diabetes from multiple demographic groups into a multiphased, mixed methods investigation of how patients use online social networking sites to communicate health information. This ongoing study consists of three phases: (1) qualitative exploration of participants’ health information communication practices (target N=36), (2) development and pilot of a survey instrument based on Phase 1 findings (target N=24), and (3) a large sample survey of patients’ health information communication practices (target N=600). This study’s findings will provide design guidance for consumer health IT supporting health information communication with patients’ social network members. As a first step, participants were asked to join our study’s Facebook group and complete a 23-question pre-study survey administered online through Qualtrics [38] assessing eligibility (ie, over 18, US citizen or residing in the United States, diagnosed with type 2 diabetes, Facebook user), demographics, Facebook use, interest in study phases, and preferred contact information. We explicitly asked about Facebook use in case participants had forwarded the survey link to individuals who were not yet members of our Facebook group. This pre-study survey primarily consists of questions previously used by research team members in other studies; new questions were developed based on the expertise of two research team members with extensive experience in survey

methodology. Results will serve as a foundation for purposive sampling for the study phases. This paper reports on recruitment activities to acquire pre-study surveys taken by the first 100 group members (the point at which maximum variance sampling for Phase 1 interviews was initiated). These activities occurred between May and June 2014.

Recruitment Procedures

Overview

Given the dearth of health sciences studies leveraging Facebook's social structures for recruitment, we sought guidance from sociological methodology. Specifically, our strategies were informed by Bhutta [39]. Bhutta's method involved reaching out to her own social network and to administrators of existing Facebook groups for baptized Catholics in the United States (Bhutta's target population). Interested individuals were directed to a Facebook group created for the study, and upon joining, were sent direct messages containing a link to the study survey.

Our strategies differed from Bhutta's in key ways. First, we reached out not only to Facebook groups, but also to pages. Second, Study 1 did not involve a study group; rather, the survey link was directly posted to relevant groups and pages. Third, in Study 2, group members were informed that the pre-study survey was an initial step in a more complex study and that they would be asked to engage in additional study-related activities such as online interviews or focus groups. Finally, to minimize potential for coercion, we did not directly advertise our study to members of our personal social networks. Below, we detail the recruitment strategy for each study.

Study 1: Consumer Health Information Technology in a Filipino Community

The methodology we developed and tested for Study 1 consisted of three parts: (1) identifying Facebook groups and pages

targeting Filipino-Americans, (2) announcing our study to administrators of identified groups and pages (ie, gatekeepers to our target population), and (3) posting a survey link directly to consenting groups and pages. Groups and pages were identified through a keyword search of the terms "Filipino America" and "Filipino USA". Pages or groups were excluded if they did not appear to be permanently based in the United States, if their members or followers did not appear to live in the United States, if their Facebook presence was inactive, if they appeared not to focus on the Filipino-American community, or if they did not allow others to initiate contact. Through this keyword search and application of exclusion criteria, we identified a total of 78 groups and 69 pages.

Administrators were contacted through Facebook's private messaging function (Figure 1). In this study, we did not pay to ensure that private messages were sent to recipients' inboxes. Consequently, some messages may have been delivered to "Other" boxes, adopted by some users to receive messages from individuals who are not among their Facebook friends. Responses were received from one group and seven page administrators. Given the low response rate, the strategy was updated to directly post our survey link to the groups and pages. Since only group members are allowed to post, we asked to join each identified group. For "Open" groups, which contain the most relaxed privacy features, membership was granted automatically. For "Closed" groups, with more extensive privacy features, permission from an administrator was required. Depending on a page's privacy settings, posts were either automatically approved or sent to the administrator for approval. In all groups and pages for which access was granted, we posted a study announcement and SurveyMonkey link (Figure 2) a total of five times over a 4-month period.

Figure 1. Message to Filipino-American group and page administrators.**Figure 2.** Post to Filipino-American groups and pages.

Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

The methodology developed and tested in this study required (1) creating a study group, (2) identifying Facebook groups and pages relevant to individuals with type 2 diabetes and racial and ethnic minorities, (3) messaging group and page administrators about our study, (4) managing the study group, and (5) messaging study group members about the pre-study survey.

As a "home" for our study on Facebook, we created a group titled, "Diabetes Management Study Community". To promote rapport, we used an image of the research team engaging in health management activities as a cover photo. We used the "About" section, visible to all Facebook users, to establish transparency of research activities. This section contained a welcome message, study description, eligibility criteria, current study activities, group rules, contact information, and information about sponsorship and institutional review board approval. Moreover, we posted files including detailed

descriptions of the study and its eligibility criteria, and a list of anticipated frequently asked questions to the group's page.

We searched Facebook to identify groups and pages for individuals with type 2 diabetes (or diabetes more generally) and for racial and ethnic minorities to promote diversity in our sample. We used keywords such as "Type 2 diabetes", "diabetes", "Hispanic", "Asian American", "Pacific Islander", "African American", "Chinese American", and "Pakistani American". It is important to note that this keyword search was

not exhaustive of all potential search terms. Groups and pages were excluded if they contained fewer than 10 members or had fewer than 10 likes, their members did not appear to live in the United States, they had been inactive for the past 6 months, or they did not allow others to initiate contact. After this keyword search and application of exclusion criteria, we contacted 122 groups and 132 pages. Additionally, we contacted 1 group and 5 pages based on referrals from those previously contacted. A breakdown by target population is shown in [Table 1](#).

Table 1. Study 2 group and page breakdown.

Target population	Groups contacted, n	Groups responding, n (%)	Pages contacted, n	Pages responding, n (%)
Type 2 diabetes	49	24 (49)	54	18 (33)
American Indian/Alaska Native	14	6 (43)	17	6 (35)
Asian	18	7 (39)	27	6 (22)
Native Hawaiian or other Pacific Islander	14	6 (43)	13	5 (39)
Black or African American	16	6 (38)	11	2 (18)
Hispanic/Latino	12	7 (58)	15	2 (13)
Total racial/ethnic	74	32 (43)	83	21 (25)
Total	123	56 (46)	137	39 (29)

Facebook's private messaging function was used to contact administrators of each identified group and page ([Figure 3](#)). Given our low response rate from administrators in Study 1, when required, we paid Facebook a nominal fee (between US \$0.19 and \$1.06) to ensure that messages were routed to people's "Inbox" as opposed to "Other" box. We paid this fee for 91 of the 123 group administrators contacted. No fee was required to send messages to pages. Follow-up messages were sent to all administrators who had not responded within 5 days. In total, administrators of 56 groups and 39 pages responded. We crafted personalized replies to each to establish trust and transparency.

The study group was managed to promote engagement without overwhelming members with excessive content. The research team posted weekly, informing members about study progress, encouraging them to recruit others ([Figure 4](#)) and responding to all questions that were not answered by other group members. All member-generated posts required approval by the research team. We requested that group members use the group's page

only to ask general questions and send questions about specific situations directly to the group's moderator. Posts containing medical advice and solicitations were not approved. Initially, the research team intended to post diabetes-specific informational resources consistent with those approved by the National Diabetes Education Program (NDEP), American Diabetes Association (ADA), or other similar institutions. An initial post by the research team, however, was met with multiple comments expressing concerns because the members didn't believe in "mainstream" guidance, but instead followed alternative methods of disease management. For the remainder of the study, the research team also adhered to the guideline of not posting messages containing medical advice.

Once our study group reached 100 members, we sent private messages to each member thanking them for joining and asking them to complete the pre-study survey. Two reminder messages were sent to all who did not complete it.

Figure 3. Message to groups and pages for Study 2.

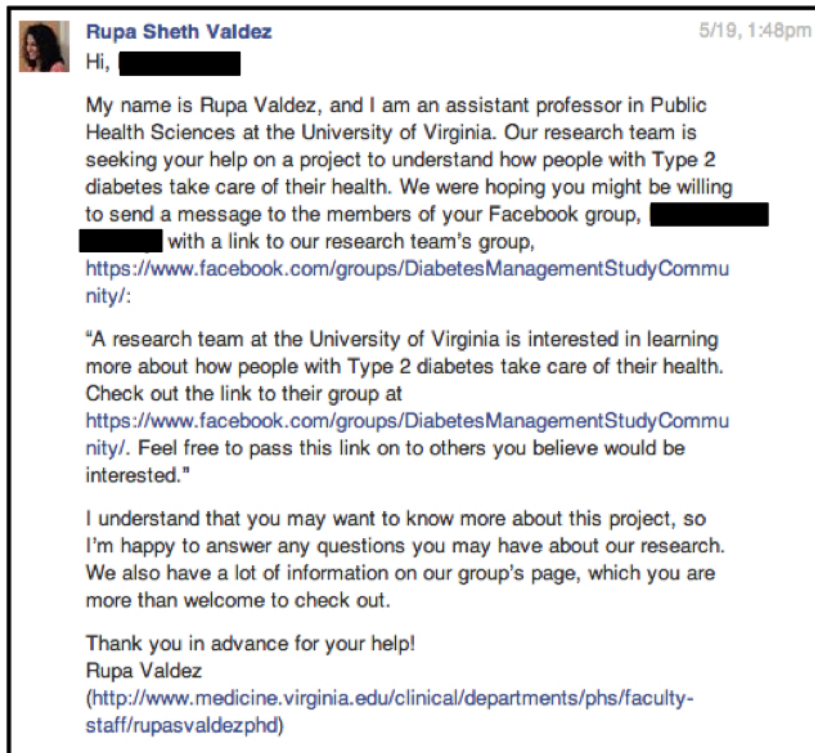
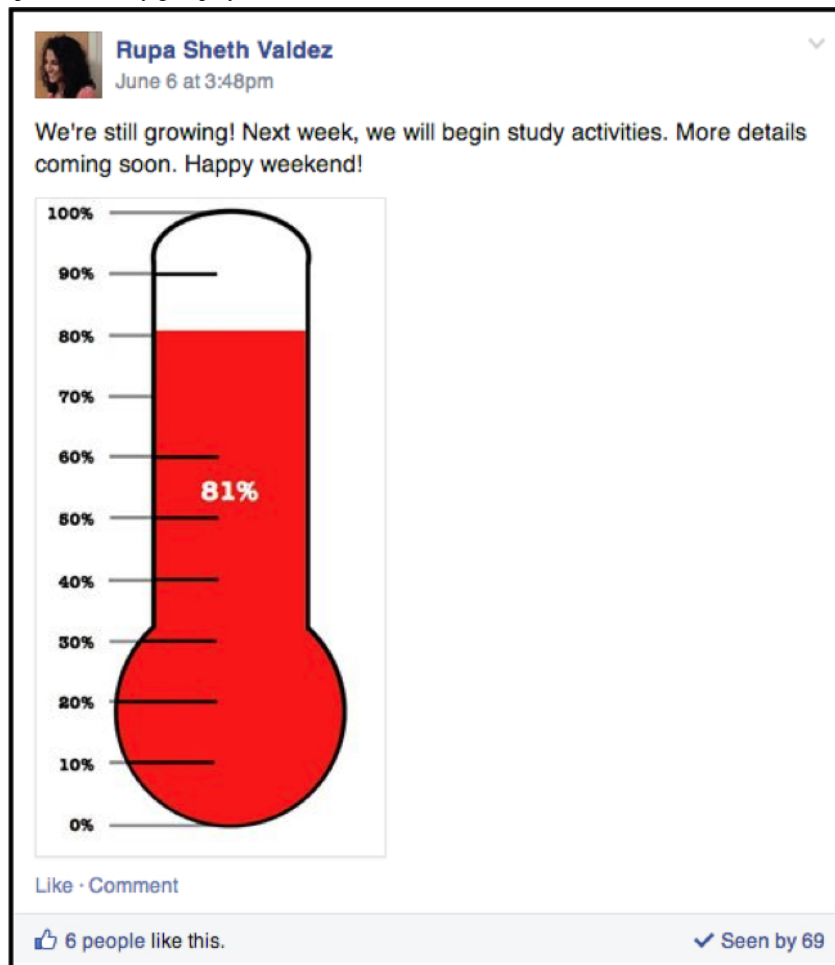


Figure 4. Example message posted to study group by research team.



Ethical Considerations

Both studies were approved by the University of Virginia's Social and Behavioral Sciences Institutional Review Board. The first screen of each survey provided details about the study's purpose and participants' rights. Advancement past the initial screen was interpreted as provision of informed consent. No incentive for survey completion was provided in either study. However, participants completing the pre-study survey for Study 2 were informed that incentives would be provided if they were recruited into Phase 1 or Phase 2.

Participant confidentiality was maintained throughout both studies. In Study 1, survey data were collected anonymously. In Study 2, survey data were not anonymous because they were used as a platform for maximum variance sampling; however, all data were securely stored on password-protected computers and files. To further ensure privacy and security of participants' information, we created a closed Facebook group, in which group membership is public, but only members can see posts and access group files. Additionally, we communicated with members about privacy and security issues to increase awareness. For example, in the "About" section, we reminded participants that their Facebook friends might be able to see what groups they joined and provided the option of communicating with us outside Facebook.

Data Analysis

All survey data were exported into SPSS 21.0 [40] for analysis. The percentage of missing data and the demographic compositions of both samples were determined using descriptive statistics. All posts and messages received from group and page administrators and study group members were loaded into QSR NVivo 10 [41] for analysis. Qualitative data from both studies were analyzed concurrently using qualitative content analysis methods [42-45] that stop at a description of the data in everyday language (as opposed to theory development). The unit of analysis was the message, and simultaneous coding was used

when units of text contained several meanings [46]. To ensure rigor, 3 investigators collaboratively coded the data. One investigator (MJT) created an initial coding framework for Study 1 and another (HKM) created an initial coding framework for Study 2. Both met together with a third (RSV) to create a revised coding framework encompassing both studies. The initial 2 investigators then recoded their respective datasets as appropriate, met again with the third investigator, and conducted a final review of the other investigator's dataset to ensure consistency.

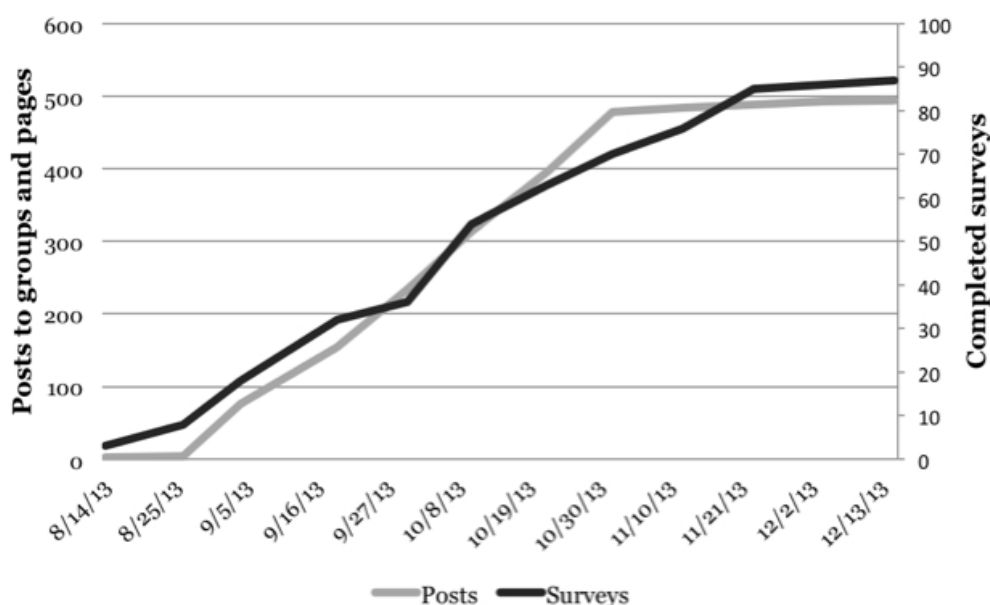
Results

Recruitment

Study 1: Consumer Health Information Technology in a Filipino Community

We successfully posted our study message to 31 of the 78 identified Facebook groups (40%) and 62 of the 69 identified Facebook pages (90%). During the study, our posts were "seen" a total of 3564 times. Facebook indicates that a post has been seen by anyone who scrolls past or follows a link directly to a post. Consequently, someone who "sees" a post may not necessarily read it. Our posts to groups received 65 "likes", and our posts to pages received 9 "likes". Additionally, our posts were shared three times, and the principal investigator received 9 friend requests from members of the target population. The survey was accessed 137 times of which 87 resulted in completed surveys. The rate of study completion was proportionate to that of posting to groups and pages (Figure 5). Approximately 6% of survey data were missing for the demographic variables and 2% for the remaining 22 close-ended variables; 91% of the respondents provided substantive feedback on at least one, 86% on at least two, and 75% on at least three of the four open-ended questions. No direct costs beyond staff labor were incurred.

Figure 5. Study 1 cumulative recruitment over time.

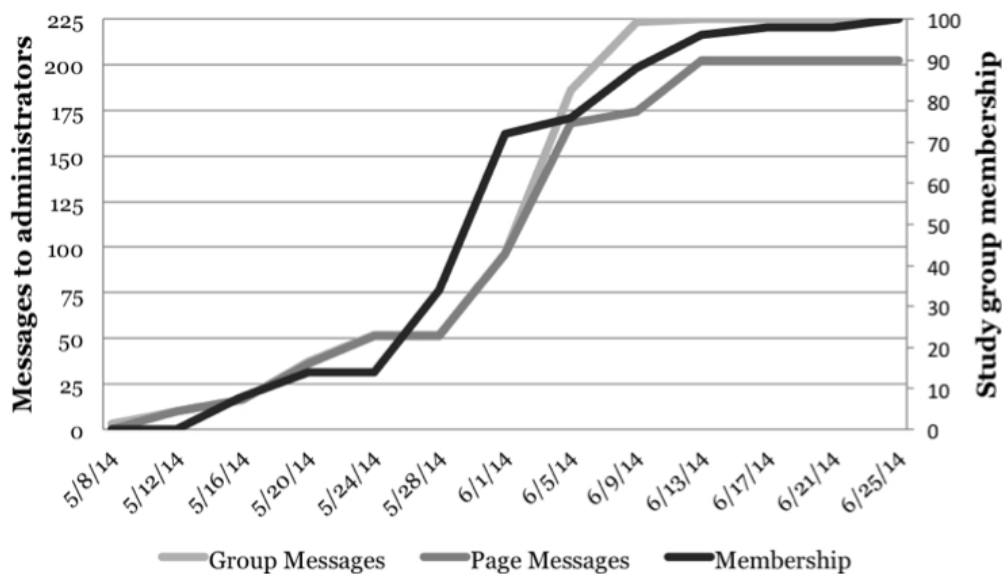


Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

We received replies from administrators of 56 of the 123 contacted groups (45%) and 39 of the 137 contacted pages (28%). Table 1 shows a breakdown by type of group and page. Eight administrators volunteered to share information about our study via Twitter, blogs, newsletters, listservs, or Google groups. Because we were not members of the groups, we were unable to determine how many Facebook users had “seen” or “liked” posts containing our study information. The principal investigator received friend requests from 3 administrators and

1 group member. Figure 6 displays numbers of messages sent to group and page administrators and membership of our study group. Membership growth seems to have followed not from the numbers of messages but rather the actions of a subset of administrators. From the 100 group members, we received 79 completed pre-study survey responses from unique individuals (after checking for duplicate Internet protocol [IP] addresses and contact information), of which 61 were from eligible individuals. Approximately 1% of survey data were missing. Total direct costs incurred (excluding investigator and staff effort) totaled US \$118.17, equaling \$1.94 per completed survey from eligible individuals.

Figure 6. Study 2 cumulative Facebook study group membership over time.



Participant Characteristics

Study 1: Consumer Health Information Technology in a Filipino Community

Demographic characteristics of the 87 respondents who completed the survey appear in Table 2. Gender was evenly distributed. All age and income brackets were represented, with

the highest number of participants between ages 30 and 49 and with a household income between US \$30,000 and \$74,999. Participants were more likely to be married, speak English at home as a primary language, and have at least an Associate’s or Bachelor’s degree. Slightly more individuals reported the Philippines, rather than the United States, as their country of birth; 93% strongly or very strongly identified as Filipino.

Table 2. Study 1 demographic characteristics (n=87).

General characteristics		n (valid %)
Gender		
	Male	40 (49)
	Female	41 (51)
Age		
	18-29	18 (22)
	30-49	26 (32)
	50-64	24 (29)
	65+	14 (17)
Education		
	Less than high school	1 (1)
	High school diploma or equivalent	2 (2)
	Some college, but no degree	11 (13)
	Associate's or Bachelor's degree	35 (42)
	Graduate degree	34 (41)
Household income, US\$		
	<\$30,000	17 (22)
	\$30,000-\$74,999	26 (34)
	\$75,000-\$99,999	13 (17)
	\$100,000+	20 (26)
Marital status		
	Married	48 (59)
	Widowed	7 (9)
	Divorced	4 (5)
	Separated	1 (1)
	Never married	22 (27)
Birth country		
	Philippines	46 (55)
	United States	36 (43)
	Other	2 (2)
Primary language at home		
	English	61 (75)
	Tagalog	13 (16)
	Other	6 (9)
Identify as Filipino		
	Very strongly	59 (70)
	Strongly	19 (23)
	Neutral	5 (6)
	Not at all strongly	1 (1)

Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

The pre-study survey was completed a total of 79 times, excluding known duplicate attempts, for a response rate of 79%. However, 18 individuals were deemed ineligible (Figure 7). Of the 61 eligible individuals (Table 3), the majority were female, between the ages of 30 and 64, and married. All but one had a

high school diploma or equivalent and 76% had at least some college education. All geographic and household income categories were represented, although only two participants reported a household income over \$150,000. Eligible individuals completing the pre-study survey were predominantly white, with moderate participation from the black/African-American and Hispanic/Latino communities. Most indicated interest in more than one study phase.

Figure 7. Reasons for Study 2 ineligibility.

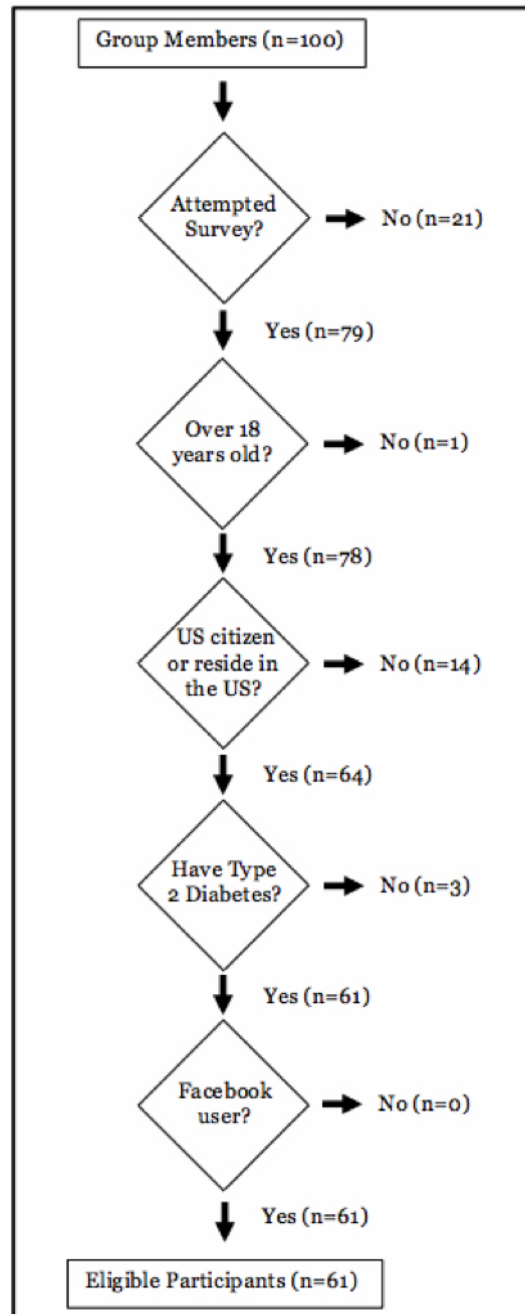


Table 3. Study 2 demographic characteristics (eligible only) (n=61).

General characteristics	n (valid %)
Gender	
Male	15 (25)
Female	46 (75)
Age	
18-29	1 (12)
30-49	22 (36)
50-64	29 (48)
65+	9 (15)
Education	
Less than high school	1 (2)
High school diploma or equivalent	11 (18)
Some college, but no degree	21 (34)
Associate's or Bachelor's degree	18 (30)
Graduate degree	10 (16)
Household income, US\$	
<\$30,000	20 (33)
\$30,000-\$74,999	22 (36)
\$75,000-\$149,999	17 (28)
\$150,000+	2 (3)
Marital status	
Married	42 (69)
Widowed	2 (3)
Divorced	11 (18)
Separated	2 (3)
Never married	4 (7)
Geographic region	
Urban	14 (23)
Suburban	20 (33)
Rural	23 (38)
Other	2 (3)
Don't know	2 (3)
Race	
White	45 (77)
Black or African American	7 (12)
American Indian/Alaskan Native	1 (2)
Asian	1 (2)
Native Hawaiian/Pacific Islander	0 (0)
Some other race	4 (7)
Ethnicity	
Hispanic or Latino	6 (10)
Not Hispanic or Latino	55 (90)
Interest in study phases	

General characteristics	n (valid %)
Phase 1	43 (70)
Phase 2	40 (66)
Phase 3	48 (79)

Content of Interaction

Study 1: Consumer Health Information Technology in a Filipino Community

We received feedback via 21 messages and 27 comments during recruitment, reflecting seven broad themes. Individuals noted their interest in providing assistance with our recruitment, confirming participation, declining participation, and providing referrals to other organizations for recruitment. They also requested clarification of multiple dimensions of our study, including our purpose, methods, and target population. A few

voiced concerns. For example, one administrator warned fellow group members about the security of their information; another expressed unease at the number of times we asked to post the recruitment message. Others provided encouragement for our research through emotionally supportive messages and instrumentally supportive messages seeking to connect us with other individuals and organizations to partner with in the future. We also received messages expressing interest in our results and us as individuals (eg, requests to be Facebook friends). Some individuals offered promotion of our study to others. Finally, we received a few messages containing solicitations. Verbatim examples are shown in [Table 4](#).

Table 4. Study 1 messages and comments received.

Theme	Study 1 message and comment quotes
Assistance	<p>“Hi Rupa,</p> <p>Thank you for reaching out to the [name of organization] facebook page for your research study. I would recommend going to our website at [url]. We have a listing of all the Filipino Organizations that belong to [name of organization] (it is an umbrella organization) with emails and contacts that may be of help to you. I unfortunately cannot post your post because our facebook page is reserved for paid partner organizations events and promotions only. I hope you find the information above helpful and good luck with your study!</p> <p>Sincerely,</p> <p>[Name] [Job title]”</p> <p>“I am Pastor of [name] Church and President of [name of organization] of [region]. I have checked out your web page at UVA and am willing to help. I will post this in both pages. Hope this helps.”</p>
Clarification	<p>“where/how are you publishing results?”</p> <p>“By the way, why did you choose to choose Filipino Community as the overall subject of this research?”</p>
Concerns	<p>“I took this survey and it has been verified as safe. No personal information will be requested. Thank you professor for your interest. For those interested using social media and other internet based applications to discuss personal health issues; be careful and use common sense. The Internet (wireless in particular) is an unsecure means of communication. Putting information about your health online may effect things such as employment and insurance.”</p> <p>“You have submitted this 4 times already...”</p>
Encouragement	<p>“Ms. Valdez, You may want to try to partner with direct service agencies who service Filipinos. Two of them are [name of organization] and [name of organization]. Both are located in [city] (within [county]) Good luck with the survey.”</p> <p>“Hello, Ms. Valdez. This is an interesting study/survey. The [name of organization] thanks you and we hope that you’ll get many participants and useful results to your survey and good luck!”</p>
Interest	<p>“No problem Rupa! I am very interested to hear about the results of your study. If you are able to would you be able to email me your findings when completed?”</p> <p>“I am the Pastor of [name of Church]. Please add me as a friend. Thank you! [url]”</p>
Promotion	<p>“Simple survey to fill out. Takes about 5-10 minutes. (Pork adobo, sisig, and chicharron...hmmm good)”</p> <p>“I encourage more people to take this survey. I already manage my healthcare online through my HMO’s website. I email my doctor and can go back and view the discussion anytime. I schedule and cancel appointments with ease online. I check pharmacies to see if my medications are available then order them and pick them up. I can view my medical test results anytime I want online. Also a medical chart ID app is like an easy access medical identification which is most valuable in case of emergencies and crisis were split second decisions are made that can save your life.”</p>
Solicitations	<p>“Hello po kabayan,</p> <p>I would like to share to you a very promising pinoy product. [name of product] - a combination of wheatgrass and guyabano. Be a dealer now and get all [name of product] products half the price (50% lifetime discount upon membership). See more of the privileges as u become part of our team! [url]”</p> <p>“Please help these families in need [url]”</p>

Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

We received 92 messages from administrators and study group members while recruiting. These reflected six of the themes seen in Study 1—*assistance, clarification, concerns, encouragement, interest, and solicitations*—and one new theme, *health information*. Individuals commented on their ability to provide recruitment assistance; however, the scope of these messages was larger than in Study 1. Specifically, we received messages stating that individuals needed time to consider participation and providing reasons for declining participation,

such as language barriers or inconsistency with the group's privacy guidelines or scope. Concerns were expressed regarding whether or not we advocated the ADA's views of diabetes management, the historical treatment of minority populations in scientific research, and the perceived vagueness of our study description. Individuals indicated interest in participating in future studies, but not in study findings or the research team. Messages related to clarification, encouragement, and solicitations paralleled those received in Study 1. A few individuals sent us messages containing health information specific to their situation. Verbatim examples are shown in [Table 5](#).

Table 5. Study 2 messages received.

Theme	Study 2 message quotes
Assistance	<p>"Hi Rupa, we thank you for reaching out and thinking of us for this cause. Unfortunately, we are unable to share your link on our page due to our community guidelines. We wish you the best of luck and thank you for your commitment toward bettering the lives of those affected by diabetes."</p> <p>"I will join the group myself and see what it is about and then I will decide if I wish to pass the information on to my other group members."</p>
Clarification	<p>"What will the information you gather go towards? New guidelines for type 2?"</p> <p>"Are you seeking this from the Arab American community or anyone can participate?"</p>
Concerns	<p>"I need to know whether your group is going to listen to concepts that may run contrary to the ADA. Many of us believe that the ADA is very narrow-minded and actually has caused more damage than helped. Need to know where you stand with traditional concepts."</p> <p>"I am an Indian advocate and it has been our cultural that these studies are never good for our culture. With broken treaties, feeding our relatives alcohol, bad government food in the name of the commodities program, isolation on the most desolate reservations and colonies. Fake blood tests that were really designed to get our DNA, trust is an issue for natives."</p>
Encouragement	<p>"Ma'am, I am not an expert on groups, but I would like to see if you would like to observe of how we run our group from the admin side. This will give you important insight on what we have found to work and what did not, how we communicate with each other, and how we run our group as tightly with less spam than most. Let me know if you would want to do this. I think it might help you to not make many of the mistakes that we did starting out."</p> <p>"Hi Rupa, Of course I will share the message! Sorry it took so long to get back to you. I was on vacation for the past two weeks and haven't done a ton of updating. Good luck in your study! Posted! J"</p>
Health Information	<p>"I have changed my eating habits and eat healthy foods now. I eat fish and chicken baked, veggies and fruits (berries, apples and citrus) occasionally bananas for the potassium. I drink a gallon of water a day and walk when possible. I take vitamin b complex, vitamin c and a multiple vitamin daily. I hope this information helps. Losing weight has helped as well. I am still working on losing more but the 40 pounds I have lost has helped a lot."</p> <p>"I have type 2 diabeteses lamtaking insulin twice aday"</p>
Interest	<p>"If you have any studies on type 1 please let me know, I would love to participate. I know [name] who works there at UVA in the [name] dept. She is a [position] with your university. I just live too far away to participate in the [name] project stuff unfortunately."</p> <p>"Thanks for the return. It is ok thanks. Well if you do get the ok to go global please keep me in your records. My wife is diabetic too by the way. May it go well and you get the results that you are searching for. Kindest regards and thanks for considering me even if unsuccessful this time round. Sincerely"</p>
Solicitations	<p>"In return could I ask that you take a look at the real diabetic reader reviews page for my book [name] [url] and send the link to people you know with Type 1 and Type 2 diabetes."</p> <p>"Hi Rupa, So sorry for the lag time. I have been on vacation and trying to get through everything. Yes, we would be happy to post this on our Facebook page. I will plan to post it tonight. I would like to ask if you would be willing to post about [page name] on your Facebook page. If so, please let me know, and I will send a post over to you. Thank you!"</p>

Study group members contributed 26 posts and 14 comments to our group's page, not including messages that remained

unapproved because they violated our community's guidelines (eg, spam, medical advice). These posts and comments reflected

three themes. Members expressed appreciation for the study and for acceptance into the Facebook group. Upon acceptance, many shared health information, including personal and family

medical histories. Others asked for and provided emotional, technical, and instrumental support to fellow community members. Verbatim examples appear in [Table 6](#).

Table 6. Study 2 posts and comments on study group's Facebook page.

Theme	Study 2 quotes of posts and comments on our Facebook group's page
Appreciation	<p>"Thank you for adding me to this group! I am a type 2 diabetic and have been since i was 16."</p> <p>"Thank you for accepting me into the group. I am a T2 for about 20 years. Have been on a pump since Dec 2009. It has changed my life. Looking forward to the study"</p>
Health Information	<p>"Diagnosed in 1992. Strong family history on my fathers side. On OmniPod since 2012. I love my pod and am convinced pumps are best. Currently in a bad slump of diabetes burn out."</p> <p>"Interesting to see where this leads. Probably borderline for 20 years.. Mild T2. Controlled by diet and exercise. Side benefit of retirement - time to exercise regularly. A1C 6.1"</p>
Support	<p>"I took the survey but haven't heard anything. I am not sure how to get to the site only got here because I got a notification! Can you tell me what to do?"</p> <p>"[name], go to Notifications on the top right and click to turn off notifications from any group you are a member of."</p>

Discussion

Principal Results

The Study 1 recruitment method yielded 87 complete survey responses, and the Study 2 method 79 complete (61 eligible) pre-study survey responses. The first method yielded a completion rate proportionate to that of the rate of posts made, whereas successes of the second method seemed to follow directly from the actions of a subset of administrators. Direct costs, excluding personnel time, were negligible, with none incurred in Study 1 and US \$118.17 in Study 2. In implementing these recruitment strategies, we received messages, posts, and comments reflecting 10 themes: appreciation, assistance, clarification, concerns, encouragement, health information, interest, promotion, solicitations, and support.

Feasibility

Leveraging Facebook for recruitment was more successful for obtaining small samples for qualitative research than large samples for quantitative research. With 87 and 61 complete responses from eligible participants in Study 1 and Study 2 respectively, our strategies did not yield sufficient participation for conducting large sample surveys, in contrast to the results obtained by Bhutta [39], whose recruitment strategies served as a basis for our methodologies. After contacting 42 groups in a 1-month period, Bhutta obtained over 4000 completed surveys from baptized Catholics in the United States. It is unclear to what this difference in sample size may be attributed. Although the prevalence of Catholicism (24.3%) [47] in the United States is greater than the prevalence of diagnosed diabetes (8.3%) [48] and the proportion of Filipinos (0.8%) [49], this difference may explain the recruitment disparity for Study 1 but is not large enough to account for the disparity in Study 2. As Bhutta conducted her study in 2008, another factor may be changing attitudes regarding privacy on Facebook and subsequent changes in use patterns [50,51]. In addition, unlike Bhutta, our methodologies did not use our personal Facebook networks for recruitment. Bhutta does not report on how participants learned about her study; however, the act of recruiting friends or friends of friends who may have greater trust in the research may have

meaningfully contributed to her study's successful recruitment outcomes.

Given the wide range of recruitment results demonstrated by researchers using Facebook advertisements, it is difficult to ascertain the effectiveness of our method in relation to this alternative. Using advertisements, health sciences investigators have reported as many as 1548 [29] and as few as zero [31] completed surveys. In terms of sample size, ours were closest to that of Lohse [27], who reported under 100 participants for each of two advertising campaigns related to nutrition. In contrast, our direct costs were substantially less than those incurred by all Facebook-based advertisement campaigns. Whereas as our strategies resulted in no direct cost (Study 1) or a cost of \$1.94 per eligible participant (Study 2), others have reported direct costs (not including personnel time) ranging from \$4.28 [29] to \$32.26 [27] per participant. Studies using Facebook advertising campaigns have reported a range of personnel time required for executing the strategy [29,32]. Our strategies required a fair amount of effort to identify Facebook groups and pages, generate personalized responses to inquiries, and moderate the study group. However, a significant portion was devoted to specifying protocols for each task. Consequently, we believe the time required to execute each strategy would be significantly reduced upon subsequent applications.

Both strategies demonstrated potential in recruiting participants for qualitative inquiry. In Study 1, a large majority of participants provided substantive responses to open-ended questions. In Study 2, we were able to recruit our target of 100 group members to serve as a sampling frame for our 36 qualitative interviews. Given that previous studies using Facebook advertisement campaigns have primarily explored their potential for quantitative research, there are no published studies that we can use to assess the effectiveness of our recruitment strategies against the use of advertisements for qualitative research.

Representativeness

Study 1: Consumer Health Information Technology in a Filipino Community

Comparisons between our study population and data from the 2012 American Community Survey (ACS) [49] were used to assess the representativeness of our sample. Whenever possible, comparisons were direct; however, for a subset of variables, direct comparisons were challenging given differences in data collection or absence of relevant data from the 2012 ACS. Our sample was reasonably representative in terms of gender, age, marital status, and birth country. In contrast to the general Filipino population living in the United States, our sample was more likely to be older than 65. This was unexpected, given that individuals over 65 are the least likely to be Facebook users [13] and that older Asian-Americans are less likely to use social media regularly [36]. Moreover, our sample contained a higher proportion of married individuals and a lower proportion of never married individuals than the US Filipino population. This may have resulted from the fact that our sample was overrepresentative of individuals with higher educational attainment, who are more likely to be married [52]. Assessing representativeness in terms of household income, language spoken at home, and identification as Filipino were infeasible given available data.

Study 2: Informing Consumer Health Information Technology Design: How Patients Use Social Networking Sites

Directly comparing our study population to the true US diabetic population on Facebook was not possible given available data. However, it is possible to estimate the proportions of the American diabetic population on Facebook within each gender and ethnic/racial category using available data and some simplifying assumptions. The estimates are based on nationwide gender and ethnic/racial population size [49], diabetes diagnosis rate for these populations [53], and estimates of Internet usage and (among Internet users) Facebook usage by these populations [13], under the simplifying assumption that gender, race/ethnicity, Facebook usage, and diabetes are independent of each other. Based on these estimates, our study was highly over representative of women, expected to comprise 51.0% of the diabetics on Facebook (women are more likely to be on Facebook, but men are more likely to have diabetes—differentials that roughly balance each other out in our calculations). To explain the gender disparity in recruitment, one would need a better understanding of how men and women use Facebook groups and pages, the basis of our recruitment. Compared with our estimated populations of diabetics on Facebook, our study was reasonably representative of whites and American Indian/Alaskan Natives, over representative of blacks/African Americans, and under representative of Hispanics/Latinos, Asians, and Native Hawaiian/Pacific Islanders. It is unclear what contributed to the underrepresentation of these latter three groups. Factors may include the activity level of our contacted Facebook groups, frequency of Facebook group use among these populations, and group differences in attitudes related to trust in scientific research. Furthermore, although we contacted groups and pages

conducted in multiple languages (eg, Spanish and English), budgetary constraints limited study participation to individuals comfortable communicating in English. We have initiated efforts to strengthen racial and ethnic representativeness of our sample by contacting additional groups and pages; this has yielded moderate improvement in the numbers of individuals identifying with currently underrepresented racial and ethnic groups.

Ethical Challenges

These studies prompted ethical deliberations. Determining whether to accept friend requests from administrators and potential or actual participants was challenging given that qualitative research requires building trust between researchers and participants while maintaining appropriate distance [54-56]. We opted not to accept friend requests, primarily due to a desire to protect the privacy of our personal lives. We considered creating alternate accounts specifically for research; however, Facebook's Statement of Rights and Responsibilities only allows one account per user [57,58]. Another option would have been to create a Facebook page to act as our public face. We decided against this option in order to facilitate peer-to-peer relationships between members of the research team and study participants. Although we can only speculate about motivations for the friend requests, it was likely that individuals wished to confirm our legitimacy and learn more about us. Consequently, instead of accepting friend requests, we actively encouraged interaction through private messages and the study's group page. Several members of the research team also expanded the information publicly available on their Facebook profiles.

Another ethical issue concerned messages from our targeted racial and ethnic groups and pages in Study 2. Although we received many supporting our recruitment efforts, we also received messages stating that our request was unrelated to the purpose of the contacted group or page. We approached groups or pages serving racial or ethnic communities of interest without excluding those unrelated to health (eg, film, engineering, photography, radio). By casting a wider net, we hoped to increase our recruitment of racial and ethnic minorities, ensuring that any recommendations for consumer health IT design were culturally relevant to multiple populations [59,60]. However, it is possible that we inadvertently alienated members of the populations we tried to engage. In future recruitment, we will explicitly state our reasons for targeting such groups (ie, to increase the representativeness of our sample).

Considerations for Future Research

Leveraging Facebook for recruitment requires building rapport with individuals online. Communication with gatekeepers and potential respondents must establish legitimacy, create trust, promote transparency, and respond effectively to concerns about participation. It is unclear what benefits and challenges researchers would encounter in repeatedly implementing this strategy, particularly with the same target population. Creating online relationships with administrators could lay a foundation for community-based research in which long-term relationships are established with community members as full partners. Conversely, frequent requests by multiple researchers may result in action against perceived spam, unless interactions are established to add greater value for participants. However,

creating greater value for participants is challenging in that participants may have varying beliefs about disease management and varying levels of comfort in sharing personal information. Broader engagement with research participants may also heighten the probability of ethical dilemmas. In particular, if the research team contains clinicians, it may be difficult to balance their multiple professional roles.

Evolving features within social media may further contribute to uncertainty about and the complexity of operationalizing this method over time and across researchers. Facebook's privacy settings are continually changing [61-63]. Researchers cannot rely on the availability of specific communication pathways with individuals to whom they are not directly connected. Similarly, algorithms for determining which posts users see remain in flux [64,65] and costs for increasing the likelihood that a post or message is seen continue to expand. Addressing

these issues effectively may require researchers to contract with specialized services that understand existing social media policies to craft protocols that best meet their needs. Moreover, a broad discussion among researchers, social media users, social media companies, and experts in research ethics is necessary to address appropriate protocols.

Conclusions

The advent of social media represents a potential solution to recruitment challenges consumer health IT researchers confront. The two studies detailed here suggest that leveraging Facebook is currently a viable means of recruitment for qualitative but not large-scale quantitative research. Given that most health-related research on Facebook has recruited through advertisements, additional research is needed to determine the long-term ethical and practical implications of adopting these alternative methods.

Acknowledgments

Study 2 is sponsored by the Agency for Healthcare Research and Quality (R03 HS22930-01). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. We would also like to thank the administrators of the Facebook groups and pages that supported our study and all of our participants. Finally, we would like to thank Anna Bickers for creating the illustrative feature image for this paper.

Conflicts of Interest

None declared.

References

1. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. *J Am Med Inform Assoc* 2014 Aug 14. [doi: [10.1136/amiajnl-2014-002826](https://doi.org/10.1136/amiajnl-2014-002826)] [Medline: [25125685](https://pubmed.ncbi.nlm.nih.gov/25125685/)]
2. Marquard JL, Zayas-Cabán T. Commercial off-the-shelf consumer health informatics interventions: recommendations for their design, evaluation and redesign. *J Am Med Inform Assoc* 2012;19(1):137-142 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2011-000338](https://doi.org/10.1136/amiajnl-2011-000338)] [Medline: [21727206](https://pubmed.ncbi.nlm.nih.gov/21727206/)]
3. Zayas-Cabán T, Dixon BE. Considerations for the design of safe and effective consumer health IT applications in the home. *Qual Saf Health Care* 2010 Oct;19 Suppl 3:i61-i67. [doi: [10.1136/qshc.2010.041897](https://doi.org/10.1136/qshc.2010.041897)] [Medline: [20959321](https://pubmed.ncbi.nlm.nih.gov/20959321/)]
4. Keselman A, Logan R, Smith CA, Leroy G, Zeng-Treitler Q. Developing informatics tools and strategies for consumer-centered health communication. *J Am Med Inform Assoc* 2008;15(4):473-483 [[FREE Full text](#)] [doi: [10.1197/jamia.M2744](https://doi.org/10.1197/jamia.M2744)] [Medline: [18436895](https://pubmed.ncbi.nlm.nih.gov/18436895/)]
5. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. *Evid Rep Technol Assess (Full Rep)* 2008 Nov(175):1-1422. [Medline: [19408968](https://pubmed.ncbi.nlm.nih.gov/19408968/)]
6. Agency for Healthcare Research and Quality. Designing consumer health IT: a guide for developers and systems designers. Rockville, Maryland; 2012. URL: <http://healthit.ahrq.gov/sites/default/files/docs/page/designing-consumer-health-it-a-guide-for-developers-and-systems-designers.pdf> [accessed 2014-10-03] [[WebCite Cache ID 6T3SZoTZk](#)]
7. Agency for Healthcare Research and Quality. Improving consumer health IT application development: lessons from other industries. Rockville, Maryland; 2011. URL: http://healthit.ahrq.gov/sites/default/files/docs/citation/background_report_final.pdf [accessed 2014-10-03] [[WebCite Cache ID 6T3ShauNQ](#)]
8. Wilson RE, Gosling SD, Graham LT. A review of Facebook research in the social sciences. *Perspect Psychol Sci* 2012 May 16;7(3):203-220. [doi: [10.1177/1745691612442904](https://doi.org/10.1177/1745691612442904)]
9. Vasalou A, Joinson AN, Courvoisier D. Cultural differences, experience with social networks and the nature of "true commitment" in Facebook. *Int J Hum Comput Stud* 2010 Oct;68(10):719-728. [doi: [10.1016/j.ijhcs.2010.06.002](https://doi.org/10.1016/j.ijhcs.2010.06.002)]
10. Hodson H. New Scientist. 2014. The secret system controlling your Facebook News Feed URL: <http://www.newscientist.com/article/mg22329804.200-the-secret-system-controlling-your-facebook-news-feed.html> [accessed 2014-10-03] [[WebCite Cache ID 6T3SmhaLy](#)]

11. The Nielsen Company. 2014. The Digital Consumer URL: <http://www.slideshare.net/tinhanhvy/the-digital-consumer-report-2014-nielsen> [accessed 2014-08-04] [WebCite Cache ID 6RZpGBj3O]
12. Facebook, Inc. Investor relations 2014 URL: <http://investor.fb.com/releasedetail.cfm?ReleaseID=842071> [accessed 2014-07-30] [WebCite Cache ID 6RSRNUqx8]
13. Duggan M, Smith A. Pew Research Center. 2013. Social media update URL: http://www.pewinternet.org/files/2013/12/PIP_Social-Networking-2013.pdf [WebCite Cache ID 6TcrQRscT]
14. Facebook, Inc. 2014. Company Info URL: <http://newsroom.fb.com/company-info/> [accessed 2014-07-30] [WebCite Cache ID 6RSRiuGeX]
15. Moran K. Forbes. 2013. The overlooked: Social media marketing for senior citizens URL: <http://www.forbes.com/sites/sap/2013/05/20/the-overlooked-social-media-marketing-for-senior-citizens/> [accessed 2014-10-03] [WebCite Cache ID 6T3T3ekOI]
16. Fox S. Pew Research Center. 2011. The social life of health information URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2014-08-04] [WebCite Cache ID 6RZpQGuec]
17. 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](https://doi.org/10.2196/jmir.1560)] [Medline: [21371990](https://pubmed.ncbi.nlm.nih.gov/21371990/)]
18. 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](https://doi.org/10.1007/s11606-010-1526-3)] [Medline: [20945113](https://pubmed.ncbi.nlm.nih.gov/20945113/)]
19. Park H, Rodgers S, Stemmler J. Health organizations' use of Facebook for health advertising and promotion. *J Interact Advert* 2011;12(1):62-77.
20. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. *Depress Anxiety* 2011 Jun;28(6):447-455 [FREE Full text] [doi: [10.1002/da.20805](https://doi.org/10.1002/da.20805)] [Medline: [21400639](https://pubmed.ncbi.nlm.nih.gov/21400639/)]
21. Whitehill JM, Brockman LN, Moreno MA. "Just talk to me": communicating with college students about depression disclosures on Facebook. *J Adolesc Health* 2013 Jan;52(1):122-127 [FREE Full text] [doi: [10.1016/j.jadohealth.2012.09.015](https://doi.org/10.1016/j.jadohealth.2012.09.015)] [Medline: [23260845](https://pubmed.ncbi.nlm.nih.gov/23260845/)]
22. Egan KG, Koff RN, Moreno MA. College students' responses to mental health status updates on Facebook. *Issues Ment Health Nurs* 2013 Jan;34(1):46-51. [doi: [10.3109/01612840.2012.718044](https://doi.org/10.3109/01612840.2012.718044)] [Medline: [23301569](https://pubmed.ncbi.nlm.nih.gov/23301569/)]
23. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME. "It's not that I don't have problems, I'm just not putting them on Facebook" challenges and opportunities in using online social networks for health. : Association for Computing Machinery; 2011 Presented at: ACM Conference on Computer Supported Cooperative Work; 2011; Hangzhou, China. [doi: [10.1145/1958824.1958876](https://doi.org/10.1145/1958824.1958876)]
24. Dillman DA, Smyth JD, Christian LM. Internet, mail and mixed-mode surveys: the tailored design method. Third Edition. New York, NY: Wiley; 2009.
25. Brooks D. New York Times. 2014. The evolution of trust URL: http://www.nytimes.com/2014/07/01/opinion/david-brooks-the-evolution-of-trust.html?_r=0 [accessed 2014-10-03] [WebCite Cache ID 6T3SBUgzc]
26. Morgan AJ, Jorm AF, Mackinnon AJ. Internet-based recruitment to a depression prevention intervention: lessons from the Mood Memos study. *J Med Internet Res* 2013;15(2):e31 [FREE Full text] [doi: [10.2196/jmir.2262](https://doi.org/10.2196/jmir.2262)] [Medline: [23403043](https://pubmed.ncbi.nlm.nih.gov/23403043/)]
27. Lohse B, Wamboldt P. Purposive facebook recruitment endows cost-effective nutrition education program evaluation. *JMIR Res Protoc* 2013;2(2):e27 [FREE Full text] [doi: [10.2196/resprot.2713](https://doi.org/10.2196/resprot.2713)] [Medline: [23948573](https://pubmed.ncbi.nlm.nih.gov/23948573/)]
28. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. *J Med Internet Res* 2012;14(1):e20 [FREE Full text] [doi: [10.2196/jmir.1978](https://doi.org/10.2196/jmir.1978)] [Medline: [22297093](https://pubmed.ncbi.nlm.nih.gov/22297093/)]
29. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *J Med Internet Res* 2012;14(1):e28 [FREE Full text] [doi: [10.2196/jmir.1878](https://doi.org/10.2196/jmir.1878)] [Medline: [22360969](https://pubmed.ncbi.nlm.nih.gov/22360969/)]
30. Arcia A. Facebook advertisements for inexpensive participant recruitment among women in early pregnancy. *Health Educ Behav* 2013 Sep 30;41(3):237-241. [doi: [10.1177/1090198113504414](https://doi.org/10.1177/1090198113504414)] [Medline: [24082026](https://pubmed.ncbi.nlm.nih.gov/24082026/)]
31. Kapp JM, Peters C, Oliver DP. Research recruitment using Facebook advertising: big potential, big challenges. *J Cancer Educ* 2013 Mar;28(1):134-137. [doi: [10.1007/s13187-012-0443-z](https://doi.org/10.1007/s13187-012-0443-z)] [Medline: [23292877](https://pubmed.ncbi.nlm.nih.gov/23292877/)]
32. Heffner JL, Wyszynski CM, Comstock B, Mercer LD, Bricker J. Overcoming recruitment challenges of web-based interventions for tobacco use: the case of web-based acceptance and commitment therapy for smoking cessation. *Addict Behav* 2013 Oct;38(10):2473-2476. [doi: [10.1016/j.addbeh.2013.05.004](https://doi.org/10.1016/j.addbeh.2013.05.004)] [Medline: [23770645](https://pubmed.ncbi.nlm.nih.gov/23770645/)]
33. Facebook, Inc. 2014. Facebook advertising guidelines URL: https://www.facebook.com/ad_guidelines.php [accessed 2014-07-30] [WebCite Cache ID 6RSS3OITs]
34. Zaid T, Burzawa J, Basen-Engquist K, Bodurka DC, Ramondetta LM, Brown J, et al. Use of social media to conduct a cross-sectional epidemiologic and quality of life survey of patients with neuroendocrine carcinoma of the cervix: a feasibility study. *Gynecol Oncol* 2014 Jan;132(1):149-153. [doi: [10.1016/j.ygyno.2013.10.015](https://doi.org/10.1016/j.ygyno.2013.10.015)] [Medline: [24145111](https://pubmed.ncbi.nlm.nih.gov/24145111/)]

35. De Choudhury M, Counts S, Horvitz EJ, Hoff A. Characterizing and predicting postpartum depression from shared Facebook data. 2014 Presented at: ACM Computer Supported Collaborative Work 2014; February 15-19, 2014; Baltimore, MD p. 626-638. [doi: [10.1145/2531602.2531675](https://doi.org/10.1145/2531602.2531675)]
36. Rainie L. Pew Internet & American Life Project. 2011. Asian-Americans and technology URL: <http://www.pewinternet.org/files/old-media/Files/Presentations/2011/Jan/2011%20-%20pdf%20-%20Asian%20Americans%20-%20DC.pdf> [accessed 2014-10-03] [WebCite Cache ID 6T3TUJIB0]
37. SurveyMonkey. surveymonkey.com. Palo Alto, CA URL: <https://www.surveymonkey.com> [accessed 2014-10-15] [WebCite Cache ID 6TUqdRsgF]
38. Qualtrics (Version 37892). Provo, UT; 2013. URL: <http://www.qualtrics.com/> [accessed 2014-10-15] [WebCite Cache ID 6TLRNLe0u]
39. Bhutta CB. Not by the book: Facebook as a sampling frame. *Sociol Method Res* 2012;41(1):57-88.
40. IBM SPSS Statistics for Windows, Version 21. Armonk, NY: IBM Corp; 2012. URL: <http://www-01.ibm.com/software/analytics/spss/> [accessed 2014-10-15] [WebCite Cache ID 6TLRXZC1G]
41. QSR International. NVivo (Version 10). 2012. URL: <http://www.qsrinternational.com/> [accessed 2014-10-15] [WebCite Cache ID 6TLRciXxV]
42. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today* 2004 Feb;24(2):105-112. [doi: [10.1016/j.nedt.2003.10.001](https://doi.org/10.1016/j.nedt.2003.10.001)] [Medline: [14769454](https://pubmed.ncbi.nlm.nih.gov/14769454/)]
43. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288. [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
44. Sandelowski M. Whatever happened to qualitative description? *Res Nurs Health* 2000 Aug;23(4):334-340. [Medline: [10940958](https://pubmed.ncbi.nlm.nih.gov/10940958/)]
45. Sandelowski M. What's in a name? Qualitative description revisited. *Res Nurs Health* 2010 Feb;33(1):77-84. [doi: [10.1002/nur.20362](https://doi.org/10.1002/nur.20362)] [Medline: [20014004](https://pubmed.ncbi.nlm.nih.gov/20014004/)]
46. Saldaña J. *The Coding Manual for Qualitative Researchers*. Thousand Oaks, CA: Sage Publications Ltd; 2009.
47. Pew Research Center. U.S. Catholics: key data from Pew Research. 2013. URL: <http://www.pewresearch.org/key-data-points/u-s-catholics-key-data-from-pew-research> [WebCite Cache ID 6TUrCK093]
48. American Diabetes Association. <https://www.diabetes.org>. 2013. Fast facts: data and statistics about diabetes URL: <http://professional.diabetes.org/admin/UserFiles/0%20-%20Sean/FastFacts%20March%202013.pdf> [accessed 2014-08-04] [WebCite Cache ID 6RZor7REb]
49. US Census Bureau. <https://www.census.gov>. 2013. 2012 American Community Survey 1-year estimates URL: <https://www.census.gov/acs> [accessed 2014-10-15] [WebCite Cache ID 6TLTTPv8y]
50. Madden M. Pew Research Center. 2012. Privacy management on social media sites URL: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Privacy_management_on_social_media_sites_022412.pdf [accessed 2014-08-04] [WebCite Cache ID 6RZp0D5NC]
51. Rainie L, Kiesler S, Kang R, Madden M. Pew Research Internet Project. 2013. Anonymity, privacy, and security online URL: <http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online/> [accessed 2014-10-03] [WebCite Cache ID 6T3TrV8Dk]
52. Pew Research Center. The decline of marriage and rise of new families. 2010. URL: <http://www.pewsocialtrends.org/2010/11/18/the-decline-of-marriage-and-rise-of-new-families/> [accessed 2014-08-04] [WebCite Cache ID 6RZp7wWzw]
53. Center for Disease Control and Prevention. National diabetes statistics report: estimates of diabetes and its burden in the United States. Atlanta, GA: Department of Health and Human Services; 2014. URL: <http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf> [accessed 2014-10-03] [WebCite Cache ID 6T3TwOzEv]
54. Guillemin M, Heggen K. Rapport and respect: negotiating ethical relations between researcher and participant. *Med Health Care Philos* 2009 Aug;12(3):291-299. [doi: [10.1007/s11019-008-9165-8](https://doi.org/10.1007/s11019-008-9165-8)] [Medline: [18853283](https://pubmed.ncbi.nlm.nih.gov/18853283/)]
55. Gordon DF. Getting close by staying distant: Fieldwork with proselytizing groups. *Qual Sociol* 1987;10(3):267-287. [doi: [10.1007/BF00988990](https://doi.org/10.1007/BF00988990)]
56. Dickson-Swift V, James EL, Kippen S, Liamputtong P. Blurring boundaries in qualitative health research on sensitive topics. *Qual Health Res* 2006 Jul;16(6):853-871. [doi: [10.1177/1049732306287526](https://doi.org/10.1177/1049732306287526)] [Medline: [16760540](https://pubmed.ncbi.nlm.nih.gov/16760540/)]
57. Facebook, Inc. 2013. Statement of rights and responsibilities URL: <https://www.facebook.com/legal/terms> [accessed 2014-07-30] [WebCite Cache ID 6RSRtDKfo]
58. Murphy J, Link MW, Childs JH, Tesfaye CL, Dean E, Stern M, et al. Social media in public opinion research. In: AAPOR Task Force on Emerging Technologies in Public Opinion Research. Deerfield, IL: American Association for Public Opinion Research; 2014.
59. Valdez RS, Gibbons MC, Siegel ER, Kukafka R, Brennan PF. Designing consumer health IT to enhance usability among different racial and ethnic groups within the United States. *Health Technol* 2012 Jul 13;2(4):225-233. [doi: [10.1007/s12553-012-0031-6](https://doi.org/10.1007/s12553-012-0031-6)]
60. Montague E, Perchonok J. Health and wellness technology use by historically underserved health consumers: systematic review. *J Med Internet Res* 2012;14(3):e78 [FREE Full text] [doi: [10.2196/jmir.2095](https://doi.org/10.2196/jmir.2095)] [Medline: [22652979](https://pubmed.ncbi.nlm.nih.gov/22652979/)]

61. McKeon M. Business Insider. 2010. The evolution of privacy on Facebook URL: <http://www.businessinsider.com/the-evolution-of-privacy-on-facebook-2010-5> [accessed 2014-09-12] [WebCite Cache ID 6SXBeaxyf]
62. Rich F. New York Magazine. 2013. Terms of engagement URL: <http://nymag.com/news/frank-rich/facebook-privacy-2013-7/> [accessed 2014-09-12] [WebCite Cache ID 6SXBIXYRV]
63. Garside J. The Guardian. 2014. Facebook bows to pressure on privacy settings for new users URL: <http://www.theguardian.com/technology/2014/may/22/facebook-privacy-settings-changes-users> [accessed 2014-09-22] [WebCite Cache ID 6SmncHZIn]
64. Carlson N. Business Insider. 2014. The hard truth about how the Facebook news feed works now URL: <http://www.businessinsider.com/how-the-facebook-news-feed-works-2014-2> [accessed 2014-09-22] [WebCite Cache ID 6Smo7zZND]
65. Facebook. Facebook for business. 2013. News feed FYI: a window into news feed URL: <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed> [accessed 2014-09-22] [WebCite Cache ID 6SmoZ7mJd]

Abbreviations

ACS: American Community Survey
ADA: American Diabetes Association
IP: Internet protocol
IT: information technology
NDEP: National Diabetes Education Program

Edited by G Eysenbach; submitted 14.08.14; peer-reviewed by B Lohse, A Arcia, K Durant; comments to author 09.09.14; revised version received 24.09.14; accepted 02.10.14; published 27.10.14.

Please cite as:

*Valdez RS, Guterbock TM, Thompson MJ, Reilly JD, Menefee HK, Bennici MS, Williams IC, Rexrode DL
Beyond Traditional Advertisements: Leveraging Facebook's Social Structures for Research Recruitment*

J Med Internet Res 2014;16(10):e243

URL: <http://www.jmir.org/2014/10/e243/>

doi: [10.2196/jmir.3786](https://doi.org/10.2196/jmir.3786)

PMID: [25348050](https://pubmed.ncbi.nlm.nih.gov/25348050/)

©Rupa S Valdez, Thomas M Guterbock, Morgan J Thompson, Jeremiah D Reilly, Hannah K Menefee, Maria S Bennici, Ishan C Williams, Deborah L Rexrode. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Tweet Content Related to Sexually Transmitted Diseases: No Joking Matter

Elia Gabarron^{1,2}, Psych; J Artur Serrano^{1,2}, PhD; Rolf Wynn^{2,3}, PhD; Annie YS Lau⁴, PhD

¹NST-Norwegian Centre for Integrated Care and Telemedicine, University Hospital North Norway, Tromsø, Norway

²Department of Clinical Medicine, Faculty of Health Sciences, The Arctic University of Norway, Tromsø, Norway

³Division of Addictions and Specialized Psychiatric Services, University Hospital North Norway, Tromsø, Norway

⁴Centre for Health Informatics, Australian Institute of Health Innovation, University of New South Wales, Sydney, Australia

Corresponding Author:

Elia Gabarron, Psych

NST-Norwegian Centre for Integrated Care and Telemedicine

University Hospital North Norway

PO Box 35

Tromsø, 9038

Norway

Phone: 47 94863460

Fax: 47 77754099

Email: elia.gabarron@telemed.no

Abstract

Background: Online social media, such as the microblogging site Twitter, have become a space for speedy exchange of information regarding sexually transmitted diseases (STDs), presenting a potential risk environment for how STDs are portrayed. Examining the types of “tweeters” (users who post messages on Twitter) and the nature of “tweet” messages is important for identifying how information related to STDs is posted in online social media.

Objective: The intent of the study was to describe the types of message emitters on Twitter in relation to two different STDs—chlamydia and human immunodeficiency virus (HIV)—as well as the nature of content tweeted, including how seriously the topic was treated.

Methods: We used the Twitter search engine to look for tweets posted worldwide from August 1-7, 2013, and from September 1-7, 2013, containing the words “chlamydia” or “HIV”, and the hashtags “#chlamydia” or “#HIV”. Tweeters were classified by two independent reviewers according to the type of avatar of the user (human, logo, or fantasy), the identification of the emitter (identifiable, semi-identifiable, or non-identifiable), and the source (private company, general media, scientific media, non-governmental, individual account, academic institution, government department, or undefined). Tweet messages were also independently classified according to their nature (serious or jokes/funny), and whether their main message was factual or of a personal nature/experience.

Results: A total of 694 tweets were posted by 426 different users during the first 7 days of August and September, containing the hashtags and/or simple words “chlamydia” and/or “HIV”. Jokes or funny tweets were more frequently posted by individual users (89%, 66/74), with a human avatar (81%, 60/74), from a non-identifiable user (72%, 53/74), and they were most frequently related to chlamydia (76%, 56/74). Serious tweets were most frequently posted by the general media (20.6%, 128/620), using a logo avatar (66.9%, 415/620), and with identifiable accounts (85.2%, 528/620). No government departments, non-governmental organizations, scientific media, or academic institutions posted a joke on STDs. A total of 104 of these analyzed tweets were re-tweeted messages, belonging to 68 unique tweets. The content was serious (99%, 67/68), factual (90%, 52/58), and about HIV (85%, 58/68).

Conclusions: Social media such as Twitter may be an important source of information regarding STDs provided that the topic is presented appropriately. Reassuringly, the study showed that almost 9/10 of tweets on STDs (chlamydia and HIV) were of serious content, and many of the tweets that were re-tweeted were facts. The jokes that were tweeted were mainly about chlamydia, and posted by non-identifiable emitters. We believe social media should be used to an even larger extent to disseminate correct information about STDs.

(*J Med Internet Res* 2014;16(10):e228) doi:[10.2196/jmir.3259](https://doi.org/10.2196/jmir.3259)

KEYWORDS

Internet; chlamydia; HIV; Twitter messaging

Introduction

In recent years, the Internet has changed the way people access general health information and make decisions about their health care [1]. The Internet has become the leading source for seeking information regarding sensitive topics, such as sexual health [2] or sexually transmitted diseases (STDs) [3].

The use of the Internet and online social media in relation to sexual health and STDs has been studied recently [4-6]. Some studies have shown that the Internet and online social media can represent a new “risk environment” in which distorted, wrong, and stigmatizing information can be published and spread rapidly [7-9], can be a space where potentially STD-infected sex partners can meet [5,10,11], and also represent a space for posting and sharing unhealthy attitudes. An example of this can be found in a study carried out on the teen dating website “Mylo1.net”, showing that among adolescents’ self-presentation, 27.7% displayed risky behavior, and 15.8% risky sexual behavior [12], with the potential effect of attracting unwanted attention from cyberbullies or sexual predators [13].

But the fact that the Internet, and specifically online social networks, are accessible to an increasing number of people, allows these powerful media tools to disseminate and inform users about evidence-based material on health [14], including sensitive topics such as sexual health and STDs [15]. Online social media are also valued as environments for their potential to engage with the general public [16,17], and particularly young people [18]. In particular, evidence suggests that certain health behaviors and sexual health behaviors might spread through social ties, of which online social networks are one example [19,20]. Recently, online social networking sites, such as Facebook or the microblogging site Twitter, have started to be used for sexual health promotion and sexual health education [10]. These online social networks may be regarded as a promising and new field for educating people about STDs.

Despite the assumed potential benefits of using online social media to promote and disseminate information on healthy sexual behaviors, there is a lack of understanding regarding how they are being used in relation to STDs. The aim of this study is to describe the nature of message emitters in the online social network Twitter on two different sexually transmitted diseases—chlamydia and human immunodeficiency virus (HIV)—the content of their tweets, and the prevalence of tweets related to unhealthy sexual health behaviors and attitudes.

Methods

Search and Data Extraction

We used the Twitter search engine to look for tweets posted worldwide in this social media, from August 1-7, 2013 and from September 1-7, 2013, containing the words “chlamydia” or “HIV”, and the hashtags “#chlamydia” or “#HIV”. Twitter was selected for being one of the fastest growing social media platforms, with roughly 500 million tweets posted every day,

and where all the posted information is fully available, even for people who have not created a Twitter account [21]. We selected these two STDs (chlamydia and HIV) because they have a common transmission mode, but with vastly different prognoses and outcomes. At the time of the study, there were two options available on the search engine to retrieve posted tweets: via keywords or via people. The search using keywords gives the option of having an overview of the people who posted in this area and the contexts in which these keywords were being used. Searches via hashtags were also conducted as they provide specific conversations focusing on these topics; hashtags are commonly used on Twitter to connect people who share a similar dialogue, and they are also used by individuals and agencies to filter tweets on a specific topic.

In regard to the months selected for examining these tweet messages, we selected August, which could be considered a common holiday month in the northern hemisphere, where the majority of Twitter users are located [22]. We chose a holiday month because people on holiday have more free time to read and post tweets and previous research has suggested there might be a higher likelihood of high-risk sexual behavior during holidays [23]. In addition, we chose September as an example of a working month.

We extracted the date of all the retrieved tweets (including both unique tweets and re-tweets), the name and avatar of emitters, and the text of the posted tweets in our analysis.

Code Categories

Emitters’ profile information (source of the emitter) and their posted messages (tone and nature), of all the downloaded tweets in English, were classified categorically by two independent reviewers (AS and EG). Any discrepancies regarding the categorization of the tweets were discussed with a third additional independent reviewer (AL) until consensus was reached. The inter-rater agreement was obtained for the three categories. A 95% confidence interval was found using the generic formula for 95% confidence intervals (estimate \pm SE 1.96). Chi-square tests were used for categorical variables. All data were analyzed with SPSS version 19 for Mac.

Tweet emitters, also called “tweeters”, were classified according to:

- Type of avatar (via image): whether the image contained at least a person, a logo image, or a fantasy image. In cases where two or more of these images were displayed at the same time in the avatar, the human image was considered. If the logo image and the fantasy image were displayed at same time in the avatar, the logo image was considered.
- Identification of the emitter (via textual profile information): identifiable (ie, the full name of the person or entity that was tweeting was clearly present), semi-identifiable (ie, identification of who was tweeting was not clearly exposed), or non-identifiable (ie, not possible to identify the person or entity responsible for the tweets).

- Source of the emitter (via textual profile information): private company, general media (newspapers, magazines, TV, and radio), scientific media (mostly scientific journals), a non-governmental organization, an individual account, an academic institution (university, college), a government department, or undefined (accounts that could not be classified in any of the previous categories).

Tweet messages were also independently classified according to:

- Tone (serious or jokes/funny): tweets were classified as jokes, according to the dictionary definition of a joke as “a thing that someone says to cause amusement or laughter, especially a story with a funny punchline” [24]. The remaining, which could not be classified as a jokes, were considered serious tweets.
- Nature: whether the main message was factual (a general topic, not related to an individual experience), or of a personal nature/experience.

Results

A total of 694 tweets (of which 104 were re-tweets) containing the hashtags and/or simple words “chlamydia” and/or “HIV”, were posted by 426 different users during the first 7 days of August and the first 7 days of September 2013. Search results are summarized in [Figure 1](#). Of those, 332 different users posted 541 tweets on HIV, and 79 different users posted 153 tweets on chlamydia. Regarding the type of avatar or image used by those 426 primary case users, 220 (51.6%) of them showed logos, 162 (38.0%) a human image, and 44 (10.3%) a fantasy avatar. In 324 cases (76.1%), the user account was considered identifiable, in 6 (1.4%) semi-identifiable, and in 96 (22.5%) non-identifiable.

The 694 tweets were classified as tweeted by individual users in 231 (33.3%) cases, the general media in 132 (19.0%) tweets, a government department in 114 (16.4%) tweets, a non-governmental organization in 90 (13.0%) tweets, scientific media in 55 (7.9%) tweets, a private company in 19 (2.7%) tweets, an academic institution in 12 (1.7%) tweets, and the nature of the emitter was undefined in 41 (5.9%) tweets.

A total of 695 tweets were downloaded, but one tweet was removed from the analysis because, although it included the word “HIV”, it was a text message written in Gaelic language and not related to STDs. Thus, the total number of tweets analyzed was 694. Regarding the categorization of the tweet messages as jokes/serious, only 14 discrepancies were found between reviewers, and the kappa value was found to be .893, almost perfect agreement, according to Landis and Koch [25]. In terms of the nature of tweets (personal experience or fact), 83 discrepancies were found in the first review round, with an inter-rater agreement of kappa=.667, considered as a substantial agreement [25]. Regarding the source of information, 145 discrepancies were found between the two reviewers, and the kappa value was .729, which represents a substantial agreement [25].

Most of the tweets posted on STDs were considered to be serious (89.3%, 620/694); however, some jokes or funny messages were found on Twitter (10.7%, 74/694), they were most frequently posted by individual users (89%, 66/74), with a human avatar (81%, 60/74) and posted by non-identifiable users (72%, 53/74). On the other side, serious tweets were most frequently posted by news organizations (20.6%, 128/620), using a logo avatar (66.9%, 415/620), and classified as identifiable accounts (85.2%, 528/620). Jokes in tweet messages were frequently related to chlamydia (76%, 56/74). No tweets of a joking nature were posted by government departments, non-governmental organizations, scientific media, or academic institutions. Tweets that contain jokes related to STDs typically convey a highly inappropriate and misguided view toward the disease, for example: “If it weren’t an STD, I would consider naming my daughter Chlamydia #pretty”, or “What’s the most positive thing in Africa?? HIV”.

[Table 1](#) describes the features of tweeters according to the nature and the tone of their tweets: funny/jokes or serious, and fact or personal experience. Of the total number of analyzed tweets, 104 were re-tweeted messages. These 104 re-tweets correspond to 68 unique tweets (seed tweets), where 4 messages were re-tweeted more than once ([Figure 2](#)). The features of the user account and content of these 68 seed tweets are summarized in [Table 2](#).

Table 1. Tone and nature of the tweets.

	Joke / Fun n=74 (10.7%)	Serious n=620 (89.3%)	Fact n=555 (80.0%)	Personal experience n=139 (20.0%)	Total N=694 (100%)
User image (avatar) ^a					
Human	60 (81.1%)	134 (21.6%)	98 (17.7%)	96 (69.1%)	194 (28.0%)
Fantasy	9 (12.2%)	71 (11.5%)	66 (11.9%)	14 (10.1%)	80 (11.5%)
Logo	5 (6.8%)	415 (66.9%)	391 (70.5%)	29 (20.9%)	420 (60.5%)
Type of account user ^a					
Identifiable	21 (28.4%)	528 (85.2%)	483 (87.0%)	66 (47.5%)	549 (79.1%)
Semi-identifiable	0 (0%)	15 (2.4%)	13 (2.3%)	2 (1.4%)	15 (2.2%)
Non-identifiable	53 (71.6%)	77 (12.4%)	59 (10.6%)	71 (51.1%)	130 (18.7%)
Tweet emitter ^a					
Individual	66 (89.2%)	165 (26.6%)	125 (22.5%)	106 (76.3%)	231 (33.3%)
General media	4 (5.4%)	128 (20.6%)	118 (21.3%)	14 (10.1%)	187 (26.9%)
Scientific media	0 (0%)	55 (8.9%)	55 (9.9%)	0 (0%)	12 (1.7%)
Government department	0 (0%)	12 (1.9%)	111 (20.0%)	3 (2.2%)	90 (13.0%)
Non-governmental	0 (0%)	90 (14.5%)	81 (14.6%)	9 (6.5%)	41 (5.9%)
Undefined	3 (4.1%)	38 (6.1%)	35 (6.3%)	6 (4.3%)	19 (2.7%)
Private company	1 (1.4%)	18 (2.9%)	18 (3.2%)	1 (0.7%)	12 (1.7%)
Academic institution	0 (0%)	12 (1.9%)	12 (2.2%)	0 (0%)	102 (14.7%)
STD ^a					
HIV	18 (24.3%)	523 (84.4%)	461 (83.1%)	80 (57.6%)	541 (78.0%)
Chlamydia	56 (75.7%)	97 (15.6%)	94 (16.9%)	59 (42.4%)	153 (22.0%)
STD search ^a					
Word HIV	18 (24.3%)	451 (72.7%)	395 (71.2%)	74 (53.2%)	469 (67.6%)
Hashtag #HIV	0 (0%)	72 (11.6%)	66 (11.9%)	6 (4.3%)	72 (10.4%)
Word chlamydia	41 (55.4%)	10 (1.6%)	8 (1.4%)	43 (90.9%)	51 (7.3%)
Hashtag #chlamydia	15 (20.3%)	87 (14.0%)	86 (15.5%)	16 (11.5%)	102 (14.7%)

^aChi-square, $P < .001$

Table 2. Features of the 68 re-tweeted messages.

Feature	Chlamydia n=10 (15%)	HIV n=58 (85%)
User image^a		
Logo	1 (10%)	0 (0%)
Human	1 (10%)	8 (14%)
Fantasy	8 (80%)	50 (86%)
Type of account user^a		
Identifiable	9 (90%)	56 (97%)
Semi-identifiable	0 (0%)	0 (0%)
Non-identifiable	1 (10%)	2 (3%)
Tweet emitter^a		
Individual	1 (10%)	8 (14%)
General media	2 (20%)	11 (19%)
Scientific media	1 (10%)	8 (14%)
Government department	6 (60%)	12 (21%)
Non-governmental	0 (0%)	15 (26%)
Undefined	0 (0%)	1 (2%)
Private company	0 (0%)	3 (5%)
Academic institution	0 (0%)	0 (0%)
Tone of the message^a		
Joke / Funny	1 (10%)	0 (0%)
Serious	9 (90%)	58 (100%)
Nature of the message^a		
Fact	9 (90%)	52 (90%)
Personal experience	1 (10%)	6 (10%)

^aChi-square, $P < .001$

Figure 1. Search and study selection process of tweets about chlamydia and HIV on Twitter.

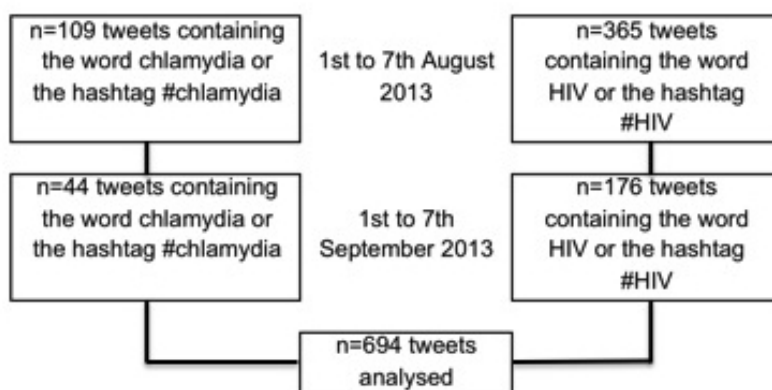


Figure 2. Examples of messages that have been re-tweeted more than once.

Seed tweet	Number of re-tweets
#Chlamydia infects more than 1,210,523! See #symptoms in men http://bit.ly/dVNbJJ	32
#Chlamydia rates highest among African American women age 15-24. Let's change this for the better http://go.usa.gov/jUsD "	4
Protect yourself: all sexually-active women age 25 & under should request a #chlamydia test each year; http://go.usa.gov/DYyR #STD"	2
Sexually-active #women age 25 & under should be screened annually for #chlamydia> Why? http://go.usa.gov/jQ4J	2

Discussion

Principal Findings

In our analysis of tweets on STDs (ie, HIV or chlamydia) posted during the first 7 days of August and September 2013, we found that nearly 9 of 10 tweets focused on HIV (620 of the total 694). This higher frequency of tweets about HIV may be explained by an increased awareness of the disease, to the detriment of chlamydia, which may be a reflection of what happens in society.

Regarding the month of the posts, almost 7 of 10 were tweeted during August (474 of the total of 694), which could be considered a holiday month. Holiday months can be seen as periods in which people have more free time to read and post tweets, but also periods in which people may have a higher rate of sexual activity, and likelihood of higher-risk sexual behavior [23]. Posting health promotion messages on Twitter during a holiday month (eg, Christmas) could potentially represent an educational or even a preventive tool in a period when there is an increased likelihood of higher-risk sexual behavior.

Some jokes or funny remarks on STDs were also posted. They were more frequently tweeted by individual users, with a human avatar, and came from a non-identifiable user (ie, very difficult to ascertain their real identity). The jokes were most frequently posted on chlamydia.

On the other hand, serious tweets on STDs were posted more frequently by users using a logo as avatar, had easily identifiable emitters, and belonged more frequently to news organizations, scientific media, and individuals. Almost all the re-tweeted messages were of serious content as well, and in these cases, the first tweet in the sequence belonged to an identifiable user too, using a fantasy avatar, and were mostly from the general media, the scientific media, government agencies, and non-governmental organizations. Regarding the messages re-tweeted, it is also interesting to stress that although most of the re-tweets were on HIV, the four messages that were re-tweeted two or more times were on chlamydia.

Tweets on personal experiences with STDs were posted more frequently by individual users, with a human avatar, and they included the word chlamydia in their tweets. Personal experience tweets were longer than tweets of a factual nature. In addition, tweets on STD-related facts were posted more frequently by

identifiable users with a logo image, and had a higher frequency of re-tweets.

Overall, sexually transmitted diseases are a worrying problem worldwide. While both diseases are sexually transmitted, HIV and chlamydia appear to be seen as very different issues in online social media. While information related to HIV stresses the disease as a serious topic (such as treatment, prevention, or stigma), chlamydia is more frequently related to joke tweets.

Behavior and Seriousness of Tweets on STDs

While online and social media communication may share many features with face-to-face communication, there are important differences [26]. Some differences may be most pronounced when the communication is anonymous. For example, self-disclosure [27] and disinhibition [28] are found to be more frequent in online communication, such as social media, where people have the option of camouflaging themselves when talking about sensitive topics, than in face-to-face mode. Anonymous communication might be beneficial for some people, for instance people who suffer from HIV or other STDs may discuss their experiences or concerns if allowed to do so anonymously. On the other hand, anonymity may be exploited negatively, for instance to make offensive statements or to deceive someone.

Our findings can be understood in at least two different ways. First, one may postulate that the majority of chlamydia joking-related messages stems primarily from anonymous tweeters [28]. In this respect, the joking behavior may be understood as resulting from the lack of social norms or group values contingent on tweeters' anonymity, where the ability to hide behind an anonymous avatar on Twitter allows one to adopt a disinhibited behavior with no responsibility or accountability.

Second, one may also see the differential treatment given to chlamydia and HIV as a result of group norms, suggesting it is acceptable to tweet chlamydia-related jokes but that tweeting HIV-related jokes is not acceptable. Following this reasoning, in situations when one can get away with anonymity, such as computer-mediated communication, the Social Identity Model of Deindividuation Effects (SIDE) [29] suggests that group norms may become even more influential, and group members tend to follow behaviors set by these existing social norms even when these behaviors are deemed unacceptable in non-anonymous settings [30,31].

This reasoning suggests that in the Twitter community, tweets related to jokes on chlamydia are more likely to be acceptable

(although most who do so prefer to do it anonymously), while jokes pertaining to HIV are unacceptable. The reason why such a distinction is made is unclear, but it may be related to the severity of the disease (ie, HIV is potentially deadly, chlamydia is not), or concerns regarding stigmatization (ie, the HIV epidemic has affected some groups of the population more than others), or other factors we are unaware of. While it is encouraging to see that only factual information or non-joking material is being re-tweeted at the time of analysis, many jokes can still be found on Twitter.

Twitter as a Channel for Sensitive Information

It is a fact that chlamydia is a less severe disease than HIV, and that if chlamydia is diagnosed early, it can easily be cured [32]. However, few people do the tests to check if they have chlamydia, and in many cases the disease may become chronic [32]. Both infections, HIV and chlamydia (and other STDs), can be spread by exactly the same sexual risk-taking behaviors [32]. We believe social media could give hard-to-reach young people unique access to high quality information about the sensitive topic of STDs. Some governmental and health organizations are already exchanging information about STDs on Twitter, such as USAID Education and the Centers for Disease Control and Prevention. However, the topic of their tweets on STDs is mainly HIV. This is possibly due to the fact that from an organizational perspective, there are agencies that have a primary mandate or focus on HIV (such as AIDS service organizations), but few have a focus specifically on chlamydia (being much more likely to have a focus on all STDs or on sexual health more broadly). Other stakeholders should pay attention to the discussion in online social media and could utilize this channel more frequently in the preventive work against STDs.

As in other online social media and on the Internet in general, consumers can find distorted information on Twitter [7,8,33,34]. Currently, there is no shortage of jokes, funny remarks, and other comments on STDs on Twitter. Although we did not classify tweets for its misguided content, some of the tweets may communicate a misguided attitude on STDs (eg, where having an STD is something to be proud of), which is not in line with the important task of reducing the spread of STDs.

Preventing and containing the spread of false information and myths on STDs should be an important item on any public health agenda. In particular, content on Twitter and other online social media outlets is available to a large number of consumers, where misguided information can spread rapidly [7-9]. Myths and misinformed attitudes that are not contained can potentially lead to dire consequences on our health attitudes, decisions, and actions [8,9,30]. Strategies for creating social awareness regarding the spread of misinformation and for finding ways to improve the quality of the information should be addressed with careful planning.

Government departments, non-governmental organizations, and academic institutions represent a more trustworthy source of information on STDs. In fact, 18% of the tweets came from government agencies (government departments and academic institutions), and 16% from non-governmental organizations, which is not a bad start but is an area that could be improved. These agencies should consider increasing their presence and visibility on social media in order to reach their target groups with high quality information, and to counteract some of the more flippant tweets that downplay the health risks of STDs.

In future research, it would be of interest to analyze other online social media, such as Facebook or YouTube, in order to examine how the topic of STDs is presented in these platforms.

Limitations

The study has several limitations. We selected two STDs with a common transmission mode, but with vastly different prognoses and outcomes (chlamydia and HIV). For the search conducted on Twitter, we only focused on tweets using the word or the hashtag “chlamydia” or “HIV”, which may have reduced the number of tweets related to the topic. Future studies should broaden the scope of keywords and hashtags, for example, including terms such as “AIDS”.

Although the search was conducted worldwide, the terms used implied that mostly tweets in English were identified.

Taking into account that we analyzed only two STDs on Twitter (HIV, chlamydia), and we used only English words, our results cannot be generalized to other online social media platforms or to other settings where English is not the primary language.

The number of Twitter followers associated with users posting on STDs was not collected. Future research should consider collecting that information, and investigate the spread rate of tweets.

Conclusions

The study showed that nearly 9 of every 10 tweets on STDs (chlamydia and HIV) were of serious content, and many of the tweets that were re-tweeted were facts. We believe this finding is reassuring as it suggests that most content on Twitter relating to STDs is of a factual and serious nature, which we hope might help in informing people about these diseases. However, many jokes could also be found, mainly about chlamydia, and these jokes and funny remarks were typically posted by non-identifiable emitters. For social media such as Twitter to be considered an important source of public health information regarding STDs, the topic needs to be presented appropriately. We believe social media plays an important role in the next generation of public health tools in disseminating correct information about STDs.

Acknowledgments

This work was supported by Helse Nord (grant numbers HST1023-11 and HST1184-14). At the time of writing, AYSL was supported by the National Health and Medical Research Council (NHMRC) Centre of Research Excellence in Informatics and E-Health (1032664).

Authors' Contributions

EG contributed to conception and design, acquisition of data, tweets review, analysis and interpretation of data, drafting, and revising the manuscript critically for important intellectual content. JAS contributed to conception and design, tweets review, and revising the manuscript critically for important intellectual content. RW contributed to conception and design, and drafting and revising the manuscript critically for important intellectual content. AYSL contributed to conception and design, tweets review, and drafting and revising the manuscript critically for important intellectual content. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Hill S, Merchant R, Ungar L. Lessons learned about public health from online crowd surveillance. *Big Data* 2013;1(3):160-167 [[FREE Full text](#)] [doi: [10.1089/big.2013.0020](https://doi.org/10.1089/big.2013.0020)] [Medline: [25045598](https://pubmed.ncbi.nlm.nih.gov/25045598/)]
2. Buhi ER, Daley EM, Fuhrmann HJ, Smith SA. An observational study of how young people search for online sexual health information. *J Am Coll Health* 2009;58(2):101-111. [doi: [10.1080/07448480903221236](https://doi.org/10.1080/07448480903221236)] [Medline: [19892646](https://pubmed.ncbi.nlm.nih.gov/19892646/)]
3. Magee JC, Bigelow L, Dehaan S, Mustanski BS. Sexual health information seeking online: a mixed-methods study among lesbian, gay, bisexual, and transgender young people. *Health Educ Behav* 2012 Jun;39(3):276-289. [doi: [10.1177/1090198111401384](https://doi.org/10.1177/1090198111401384)] [Medline: [21490310](https://pubmed.ncbi.nlm.nih.gov/21490310/)]
4. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. *Cochrane Database Syst Rev* 2010(9):CD006483. [doi: [10.1002/14651858.CD006483.pub2](https://doi.org/10.1002/14651858.CD006483.pub2)] [Medline: [20824850](https://pubmed.ncbi.nlm.nih.gov/20824850/)]
5. Fairley CK. Using information technology to control STIs. *Sex Transm Infect* 2011 Dec;87 Suppl 2:ii25-ii27 [[FREE Full text](#)] [doi: [10.1136/sti.2010.048330](https://doi.org/10.1136/sti.2010.048330)] [Medline: [22110149](https://pubmed.ncbi.nlm.nih.gov/22110149/)]
6. Labacher L, Mitchell C. Talk or text to tell? How young adults in Canada and South Africa prefer to receive STI results, counseling, and treatment updates in a wireless world. *J Health Commun* 2013;18(12):1465-1476. [doi: [10.1080/10810730.2013.798379](https://doi.org/10.1080/10810730.2013.798379)] [Medline: [24015829](https://pubmed.ncbi.nlm.nih.gov/24015829/)]
7. Lau AY, Gabarron E, Fernandez-Luque L, Armayones M. Social media in health--what are the safety concerns for health consumers? *HIM J* 2012;41(2):30-35. [Medline: [23705132](https://pubmed.ncbi.nlm.nih.gov/23705132/)]
8. Lau AY, Coiera EW. Impact of web searching and social feedback on consumer decision making: a prospective online experiment. *J Med Internet Res* 2008;10(1):e2 [[FREE Full text](#)] [doi: [10.2196/jmir.963](https://doi.org/10.2196/jmir.963)] [Medline: [18244893](https://pubmed.ncbi.nlm.nih.gov/18244893/)]
9. Lau AY, Kwok TM, Coiera E. How online crowds influence the way individual consumers answer health questions: an online prospective study. *Appl Clin Inform* 2011;2(2):177-189 [[FREE Full text](#)] [doi: [10.4338/ACI-2011-01-RA-0006](https://doi.org/10.4338/ACI-2011-01-RA-0006)] [Medline: [23616869](https://pubmed.ncbi.nlm.nih.gov/23616869/)]
10. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. *BMC Public Health* 2011;11:583 [[FREE Full text](#)] [doi: [10.1186/1471-2458-11-583](https://doi.org/10.1186/1471-2458-11-583)] [Medline: [21777470](https://pubmed.ncbi.nlm.nih.gov/21777470/)]
11. Swendeman D, Rotheram-Borus MJ. Innovation in sexually transmitted disease and HIV prevention: internet and mobile phone delivery vehicles for global diffusion. *Curr Opin Psychiatry* 2010 Mar;23(2):139-144 [[FREE Full text](#)] [doi: [10.1097/YCO.0b013e328336656a](https://doi.org/10.1097/YCO.0b013e328336656a)] [Medline: [20087189](https://pubmed.ncbi.nlm.nih.gov/20087189/)]
12. Pujazon-Zazik MA, Manasse SM, Orrell-Valente JK. Adolescents' self-presentation on a teen dating web site: a risk-content analysis. *J Adolesc Health* 2012 May;50(5):517-520. [doi: [10.1016/j.jadohealth.2011.11.015](https://doi.org/10.1016/j.jadohealth.2011.11.015)] [Medline: [22525117](https://pubmed.ncbi.nlm.nih.gov/22525117/)]
13. Pujazon-Zazik M, Park MJ. To tweet, or not to tweet: gender differences and potential positive and negative health outcomes of adolescents' social internet use. *Am J Mens Health* 2010 Mar;4(1):77-85. [doi: [10.1177/1557988309360819](https://doi.org/10.1177/1557988309360819)] [Medline: [20164062](https://pubmed.ncbi.nlm.nih.gov/20164062/)]
14. Gabarrón E, Fernández-Luque L. eHealth and online videos for health promotion. *Gac Sanit* 2012;26(3):197-200 [[FREE Full text](#)] [doi: [10.1016/j.gaceta.2012.03.004](https://doi.org/10.1016/j.gaceta.2012.03.004)] [Medline: [22554457](https://pubmed.ncbi.nlm.nih.gov/22554457/)]
15. Lengen C, Jäger S, Kistemann T. The knowledge, education and behaviour of young people with regard to Chlamydia trachomatis in Aarhus, Denmark and Bonn, Germany: do prevention concepts matter? *Soc Sci Med* 2010 Jun;70(11):1789-1798. [doi: [10.1016/j.socscimed.2010.01.048](https://doi.org/10.1016/j.socscimed.2010.01.048)] [Medline: [20307923](https://pubmed.ncbi.nlm.nih.gov/20307923/)]
16. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. *Health Promot Pract* 2012 Mar;13(2):159-164. [doi: [10.1177/1524839911433467](https://doi.org/10.1177/1524839911433467)] [Medline: [22382491](https://pubmed.ncbi.nlm.nih.gov/22382491/)]

17. Neiger BL, Thackeray R, Burton SH, Giraud-Carrier CG, Fagen MC. Evaluating social media's capacity to develop engaged audiences in health promotion settings: use of Twitter metrics as a case study. *Health Promot Pract* 2013 Mar;14(2):157-162. [doi: [10.1177/1524839912469378](https://doi.org/10.1177/1524839912469378)] [Medline: [23271716](https://pubmed.ncbi.nlm.nih.gov/23271716/)]
18. O'Keeffe GS, Clarke-Pearson K, Council on Communications Media. The impact of social media on children, adolescents, and families. *Pediatrics* 2011 Apr;127(4):800-804 [FREE Full text] [doi: [10.1542/peds.2011-0054](https://doi.org/10.1542/peds.2011-0054)] [Medline: [21444588](https://pubmed.ncbi.nlm.nih.gov/21444588/)]
19. Johnsen JA, Vambheim SM, Wynn R, Wangberg SC. Language of motivation and emotion in an internet support group for smoking cessation: explorative use of automated content analysis to measure regulatory focus. *Psychol Res Behav Manag* 2014;7:19-29 [FREE Full text] [doi: [10.2147/PRBM.S54947](https://doi.org/10.2147/PRBM.S54947)] [Medline: [24470780](https://pubmed.ncbi.nlm.nih.gov/24470780/)]
20. Vambheim SM, Wangberg SC, Johnsen JA, Wynn R. Language use in an internet support group for smoking cessation: development of sense of community. *Inform Health Soc Care* 2013 Jan;38(1):67-78. [doi: [10.3109/17538157.2012.710685](https://doi.org/10.3109/17538157.2012.710685)] [Medline: [22958060](https://pubmed.ncbi.nlm.nih.gov/22958060/)]
21. Holt R. The Telegraph. 2013 Mar 21. Twitter in numbers URL: <http://www.telegraph.co.uk/technology/twitter/9945505/Twitter-in-numbers.html> [accessed 2014-05-25] [WebCite Cache ID 6PpXTOLkE]
22. Beevolve. 2012. An exhaustive study of Twitter users across the world URL: <http://www.beevolve.com/twitter-statistics/> [accessed 2014-01-14] [WebCite Cache ID 6Mch04RI5160httpwwwbeevolvecomtwitterstatistics]
23. Rogstad KE. Sex, sun, sea, and STIs: sexually transmitted infections acquired on holiday. *BMJ* 2004 Jul 24;329(7459):214-217 [FREE Full text] [doi: [10.1136/bmj.329.7459.214](https://doi.org/10.1136/bmj.329.7459.214)] [Medline: [15271833](https://pubmed.ncbi.nlm.nih.gov/15271833/)]
24. Oxford Dictionaries. England: Oxford University Press; 2014. Definition of joke in English URL: <http://www.oxforddictionaries.com/definition/english/joke> [accessed 2014-09-11] [WebCite Cache ID 6SVMtsqUC]
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
26. Joinson AN. Understanding the psychology of Internet behaviour: virtual worlds, real lives. New York: Palgrave Macmillan; 2003.
27. Joinson AN. Knowing me, knowing you: reciprocal self-disclosure in Internet-based surveys. *Cyberpsychol Behav* 2001 Oct;4(5):587-591. [Medline: [11725651](https://pubmed.ncbi.nlm.nih.gov/11725651/)]
28. Joinson AN. Social desirability, anonymity, and Internet-based questionnaires. *Behav Res Methods Instrum Comput* 1999 Aug;31(3):433-438. [Medline: [10502866](https://pubmed.ncbi.nlm.nih.gov/10502866/)]
29. Reicher SD, Spears R, Postmes T. A social identity model of deindividuation phenomena. *Eur Rev Soc Psychol* 1995 Jan;6(1):161-198. [doi: [10.1080/14792779443000049](https://doi.org/10.1080/14792779443000049)]
30. Postmes T, Spears R, Sakhel K, de Groot D. Social influence in computer-mediated communication: the effects of anonymity on group behavior. *Personality and Social Psychology Bulletin* 2001 Oct 01;27(10):1243-1254. [doi: [10.1177/01461672012710001](https://doi.org/10.1177/01461672012710001)]
31. Sassenberg K, Boos M. Attitude change in computer-mediated communication: effects of anonymity and category norms. *Gr Proc Intergr Rel* 2003;6(4):405-422. [doi: [10.1177/13684302030064006](https://doi.org/10.1177/13684302030064006)]
32. World Health Organization. 2013. Sexually transmitted infections (STIs) URL: <http://www.who.int/mediacentre/factsheets/fs110/en/> [accessed 2014-04-30] [WebCite Cache ID 6PDv3CdRE]
33. Lau AY, Siek KA, Fernandez-Luque L, Tange H, Chhanabhai P, Li SY, et al. The role of social media for patients and consumer health. Contribution of the IMIA Consumer Health Informatics Working Group. *Yearb Med Inform* 2011;6(1):131-138. [Medline: [21938338](https://pubmed.ncbi.nlm.nih.gov/21938338/)]
34. Coiera E, Aarts J, Kulikowski C. The dangerous decade. *J Am Med Inform Assoc* 2012;19(1):2-5 [FREE Full text] [doi: [10.1136/amiajnl-2011-000674](https://doi.org/10.1136/amiajnl-2011-000674)] [Medline: [22116642](https://pubmed.ncbi.nlm.nih.gov/22116642/)]

Abbreviations

HIV: human immunodeficiency virus

STD: sexually transmitted disease

Edited by G Eysenbach; submitted 19.01.14; peer-reviewed by T Veinot, M Gilbert, D McDonnell, S Patolia; comments to author 26.04.14; revised version received 29.05.14; accepted 21.08.14; published 06.10.14.

Please cite as:

Gabarron E, Serrano JA, Wynn R, Lau AYS

Tweet Content Related to Sexually Transmitted Diseases: No Joking Matter

J Med Internet Res 2014;16(10):e228

URL: <http://www.jmir.org/2014/10/e228/>

doi: [10.2196/jmir.3259](https://doi.org/10.2196/jmir.3259)

PMID: [25289463](https://pubmed.ncbi.nlm.nih.gov/25289463/)

©Elia Gabarron, J Artur Serrano, Rolf Wynn, Annie YS Lau. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Tweeting for and Against Public Health Policy: Response to the Chicago Department of Public Health's Electronic Cigarette Twitter Campaign

Jenine K Harris¹, PhD; Sarah Moreland-Russell¹, PhD; Bechara Choucair², MD; Raed Mansour², MS; Mackenzie Staub¹; Kendall Simmons¹

¹Washington University in St. Louis, St. Louis, MO, United States

²Chicago Department of Public Health, Chicago, IL, United States

Corresponding Author:

Jenine K Harris, PhD

Washington University in St. Louis

One Brookings Drive

St. Louis, MO,

United States

Phone: 1 314 935 3522

Fax: 1 314 935 3756

Email: harrisj@wustl.edu

Abstract

Background: In January 2014, the Chicago City Council scheduled a vote on local regulation of electronic cigarettes as tobacco products. One week prior to the vote, the Chicago Department of Public Health (CDPH) released a series of messages about electronic cigarettes (e-cigarettes) through its Twitter account. Shortly after the messages, or tweets, were released, the department's Twitter account became the target of a "Twitter bomb" by Twitter users sending more than 600 tweets in one week against the proposed regulation.

Objective: The purpose of our study was to examine the messages and tweet patterns in the social media response to the CDPH e-cigarette campaign.

Methods: We collected all tweets mentioning the CDPH in the week between the e-cigarette campaign and the vote on the new local e-cigarette policy. We conducted a content analysis of the tweets, used descriptive statistics to examine characteristics of involved Twitter users, and used network visualization and descriptive statistics to identify Twitter users prominent in the conversation.

Results: Of the 683 tweets mentioning CDPH during the week, 609 (89.2%) were anti-policy. More than half of anti-policy tweets were about use of electronic cigarettes for cessation as a healthier alternative to combustible cigarettes (358/609, 58.8%). Just over one-third of anti-policy tweets asserted that the health department was lying or disseminating propaganda (224/609, 36.8%). Approximately 14% (96/683, 14.1%) of the tweets used an account or included elements consistent with "astroturfing"—a strategy employed to promote a false sense of consensus around an idea. Few Twitter users were from the Chicago area; Twitter users from Chicago were significantly more likely than expected to tweet in support of the policy.

Conclusions: Our findings may assist public health organizations to anticipate, recognize, and respond to coordinated social media campaigns.

(*J Med Internet Res* 2014;16(10):e238) doi:[10.2196/jmir.3622](https://doi.org/10.2196/jmir.3622)

KEYWORDS

Twitter; health departments; e-cigarette

Introduction

The use of electronic cigarettes (e-cigarettes) is growing rapidly in the United States. In a single year, e-cigarette use in the

United States doubled for adults [1] and middle and high school students [2]. Adults and youth believe e-cigarettes are less harmful than combustible cigarettes [3,4] and nearly half of Americans believe e-cigarettes should be made available for

smoking cessation [4]. Evidence related to e-cigarette effectiveness in promoting smoking cessation is mixed; studies have found increased readiness and confidence to quit after 1 week of e-cigarette use [5] and success in reducing or eliminating smoking in smokers [6,7] while other research found cessation rates were not significantly higher when e-cigarettes were used compared to other cessation aids [8]. While health risks associated with e-cigarette use are not well understood [9,10], a limited number of studies examining the chemical composition of what is inhaled by an e-cigarette user suggests surveillance and additional research is warranted [11-15].

Emerging evidence suggests that e-cigarette use is associated with smoking among adolescents and may encourage smoking initiation [16]. In light of falling smoking rates in the United States, established and emerging tobacco companies are investing in e-cigarettes and other smokeless tobacco to recruit new and retain existing customers [8]. E-cigarette marketing expenditures have increased rapidly in recent years in the United States [17], as has exposure of youth and young adults to e-cigarette advertising [18]. E-cigarettes are marketed on social media [4,19,20], and come in a high-tech format and sweet flavors like cotton candy, which may attract younger consumers [2,4,21,22]. Use of e-cigarettes is increasing in youth and young adults, including in youth and young adults who do not smoke traditional cigarettes [23].

On April 24, 2014, the Food and Drug Administration (FDA) proposed a new rule to extend existing tobacco authority to e-cigarettes and other tobacco products [24]. However, current FDA authority prohibits advertising of e-cigarettes as a therapeutic device for smoking cessation, but does not regulate e-cigarettes as a tobacco product [9]. As of January 3, 2014, three state laws and 108 local laws restricted e-cigarette use in 100% smoke-free venues and nine states restricted use in other venues [25]. On January 15th, 2014, the Chicago City Council voted to regulate e-cigarettes as tobacco products [26]. The regulation defines e-cigarettes as tobacco products and applies Chicago's tobacco control laws to e-cigarettes (record #SO2013-6160 [27]).

Informing and educating constituents about health is one of 10 essential services provided by local health departments (LHDs) in the United States [28]. Communication with constituents about health and health risks is required of LHDs for accreditation [29]. In past research, 61% of LHDs met standards for informing and educating constituents [30], suggesting a gap between current practices and best practice for communicating with constituents. LHDs have begun to adopt Twitter and other social media for communication, which may help fill the gap [31,32]. Currently, LHDs use social media primarily to broadcast information rather than interact with constituents [33-35].

One week prior to the e-cigarette policy vote, the Chicago Department of Public Health (CDPH) used Twitter to disseminate a series of tweets about e-cigarettes. Shortly after releasing the tweets, the CDPH Twitter feed was "Twitter bombed" by Twitter users tweeting against the proposed policy change. Twitter bombing is a social media strategy designed to fill Twitter feeds with a specific message in order to "establish a false sense of group consensus about a particular idea" and to

make the message a trending topic on Twitter [36]. In politics, this strategy has been called "astroturfing", which is a movement that appears to be grassroots, but is supported by a corporation, industry trade association, political interest group, or public relations firm [37-40]. The tobacco industry has historically used astroturfing by working with third-party allies [37,38] and including citizen-driven groups such as the smokers' rights movement to unite and oppose tobacco policies [37,40,41].

Astroturfing through social media has been used with some success in elections for and against candidates and issues [36,42,43]. To better understand how social media was used to oppose the Chicago e-cigarette policy, we examined patterns of Twitter use, connections among Twitter users, and content of tweets sent during the week between the campaign and the vote. To our knowledge, this is the first analysis of a Twitter bombing of a public health organization or topic. This information will aid LHDs and other organizations in anticipating, recognizing, and responding to coordinated social media strategies.

Methods

Study Context

On January 8, 2014, CDPH conducted a 1-day Twitter campaign about e-cigarettes (Table 1). CDPH Twitter campaigns are typically short, given the many priorities of the department and the varied interests of constituents. CDPH used the hashtag #ECigTruths to facilitate engagement with the tweets. Hashtags (#) are metadata embedded in tweets allowing users to click on the hashtagged word and see all tweets using the same hashtag. Hashtags facilitate group formation around specific topics or events [44] and are positively associated with retweeting [45,46]. The CDPH also directly invited engagement by ending the first e-cigarette tweet with, "Let's talk about it!"

Throughout January 8th, 12 pro-policy and 11 anti-policy tweets were directed to the CDPH Twitter feed by including the CDPH Twitter handle @ChiPublicHealth in the tweet. Including a Twitter handle in a tweet is called "mentioning" [47]. If the mention is *first* in the tweet, the tweet shows up in the Twitter timeline of those who follow both the tweeter and receiver of the mention; if the mention is later in the tweet, it is treated like a regular tweet and shows up for followers of the tweeter. Both types of mentions are visible to the owners of @ChiPublicHealth in the mentions section of the account; if the mention is in a public tweet, the tweet is also publicly available.

At 12:46am January 9th, Twitter user @A tweeted the following, "We need to twitter bomb the hell out if [sic] @chiPublicHealth, spreading nothing but lies #kcavo #vaping...[URL]." The tweet was visible to the followers of @A and Twitter users following #kcavo and/or #vaping. The hashtag #kcavo is short for a saying among e-cigarette users, or *vapers*, "keep calm and vape on". The Twitter profile of @A connects to a professional website that links to pro-vaping advocacy groups, conferences, and an e-cigarette-related business. Over the next 6 days more than 600 tweets were sent opposing the e-cigarette policy and including @ChiPublicHealth.

Table 1. E-cigarette tweets sent by the Chicago Department of Public Health (@ChiPublicHealth) on January 8, 2014^a.

Time	Tweet
10:05am	#ECigs look like, are labeled & contain nicotine like cigarettes. They should be regulated as such. Let's talk about it! #ecigtruths
10:10am	#ECigs come in cotton candy, bubble gum & gummy bear flavors - clearly meant for children [URL] #ecigtruths
10:14am	The "water vapor" from #ECigs contains benzene, nickel, tin, arsenic, formaldehyde & acrolein #ecigtruths [URL]
10:18am	Percentage of middle school and high school students who used e-cigarettes DOUBLED from 2011 to 2012. They must be regulated. #ecigtruths
10:22am	Electronic cigarettes contain a dangerous, addictive drug & should be regulated like other nicotine products #ecigtruths
10:25am	We have a duty to protect our children from ever picking up a nicotine habit #ecigtruths [URL]
10:28am	Youth are particularly susceptible to behavioral advertising [URL] #ecigtruths
10:30am	We do not want to create a new generation of nicotine-addicted residents. It's time to regulate #ecigtruths [URL]
10:33am	In Chicago, smoking rates are lower than ever. Let's not reverse decades of life-saving progress #ecigtruths
10:38am	"9 Terribly Disturbing Things About Electronic Cigarettes" [URL] via @HuffPostBiz #ecigtruths
11:34am	Electronic cigs contain a dangerous, addictive drug & should be regulated like other nicotine products #ecigtruths
12:37pm	Safe? #ecigtruths [URL]

^aSee [Multimedia Appendix 1](#) for table with URLs included.

Data Collection and Management

Using the twitterR package [48] for R version 2.15.2, we collected 684 tweets mentioning @ChiPublicHealth between January 8, 2014 and January 15, 2014. With each tweet, we collected the date and time it was sent, whether it was an original tweet or a native retweet, who sent the tweet, and who composed the tweet. Retweeting is the sharing of tweets originally composed by a different Twitter user. A native retweet is a retweet using an automated function in Twitter; when this function is used, metadata is added to the tweet marking it as a retweet and attaching the handle of the original tweeter. Using NodeXL [49,50], we collected information about each tweeter including number of followers, following, and total tweets sent. Finally, the CDPH tracked social media activity about the policy from non-CDPH sources during the week.

In reviewing the data, we noticed duplicate messages not marked as retweets. We identified these as non-native retweets (retweeted manually rather than using the automated function), coded them as retweets, and attributed each to its original tweeter. To examine how messages spread, retweet network data was visualized and analyzed using Pajek 3.10 [51,52].

Coding Tweet Content

To examine tweet content, four of the authors read the tweets independently and identified themes. The authors then met to discuss and developed a common set of themes: safety, lies/propaganda, science, flavors, regulation, and issue salience (Table 2). Tweets were assigned as many themes as relevant. We also coded each tweet as in support, opposed to, or unclear on regulating e-cigarettes. Each tweet was coded by two authors. The team met to discuss disagreements and agreed on final classifications. One tweet was not about e-cigarettes and was dropped.

Data Analysis

Patterns of Twitter use can indicate whether a conversation developed organically or through strategies consistent with astroturfing. Characteristics of accounts used for astroturfing include (1) a high following-to-follower ratio, which indicates that the user follows many others but is not followed [53,54], (2) the use of a combination of letters and numbers in a username [54], and (3) account age, with newer accounts more likely to belong to spammers [55]. We considered following-to-follower ratios more than two standard deviations above the mean ($z \geq 1.96$) to be high. Tweeters with two or more of these three characteristics we considered likely to be involved in astroturfing. Tweet characteristics can also aid in identifying messages consistent with astroturfing. For example, sending multiple copies of the same tweet that include different shortened versions of the same URL is a strategy used to avoid being identified by Twitter as a spammer [53].

We used descriptive statistics to examine Twitter user characteristics, message content, and Twitter use patterns. We used network analysis and visualization to examine patterns of retweeting. Specifically, we examined the network consisting of Twitter users with retweeting relationships connecting them. A retweet network is directed, with relationship direction representing information flow. For example, if network member B retweeted a message from network member A, the relationship would be depicted A→B since information is traveling from A to B. In a directed network, outdegree centrality measures how many outgoing ties a network member has. In this case, A→B indicates an outdegree of 1 for A, who was retweeted one time by B. Highly central network members aid in the spread or dissemination of information [56]. Tie strength is also influential in dissemination. In a directed network, asymmetric ties are typically weaker relationships, which facilitate dissemination. Mutual ties are stronger and hinder dissemination since both members in the relationship likely already have the same information [57]. To understand the network members and ties

facilitating dissemination policy tweets, we examined (1) network members with high outdegree centrality, and (2) the distribution of asymmetric and mutual ties across the network.

Results

Summary

Between January 8, 2014 and January 15, 2014, 306 Twitter users mentioning @ChiPublicHealth each sent an average of 0.83 (SD 1.71) original tweets and 1.40 (SD 1.87) retweets. On average, each user had 776 followers (range 0-38,662) and had tweeted 3063 times (range 2-143,118) during the time their account was open. A total of 35 (11.4%) of the 306 user profiles reported their location as the Chicago area and 13 (4.2%) reported Illinois, while 128 (41.7%) were blank or unknown (eg, “a galaxy far far away”), 83 (27.0%) were from another US state, and 48 (15.6%) were from outside the United States. A few accounts had no followers; to compute the following-to-follower ratio, we added one follower to these accounts. The average ratio of following-to-followers was 1.98 (range 0-19). The oldest Twitter account was opened on March 11, 2007; the newest was opened January 14, 2014. Of the 683 tweets, 258 (37.8%) started with @ChiPublicHealth. Finally, #ECigTruths was included in 174 anti-policy and 43 pro-policy tweets.

A total of 105 Twitter user accounts had one or more of the three indicators of astroturfing; 17 Twitter users had significantly higher than average following-to-follower ratios ($z \geq 1.96$), seven accounts were opened during the time of the Twitter bombing, and 94 accounts had letters and numbers in the handle. In total, 11 Twitter users had two or three characteristics and were therefore considered likely astroturf accounts.

Tweet Content

Of the 683 tweets, 51 (7.5%) were in support of e-cigarette regulation (pro-policy), while 609 of the 683 tweets (89.2%) were opposed (anti-policy), and 23 of the 683 tweets (3.4%) were coded as unable to tell. Tweet sentiment was significantly associated with location ($\chi^2_4=95.9$; $P<.001$), with standardized residuals indicating Chicago residents were significantly more likely than expected to send a pro-policy tweet. Specifically, 37 Twitter users sent at least one pro-policy tweet; 21 (56.8%) of these were located in Chicago, while 4 were elsewhere in Illinois, 10 were in unknown locations, and two were in other states.

Pro-policy tweets were most likely to focus on regulation ($n=44$) and science ($n=19$), while anti-policy tweets were most likely to focus on safety ($n=358$) and lies/propaganda ($n=224$). Tweets classified as lies/propaganda included several types of arguments diverting attention from the original e-cigarette messages rather than addressing message substance. For example, a few lies/propaganda tweets resorted to name calling (eg, “@ChiPublicHealth are a bunch of IGNORANT LIARS”) while many focused on discrediting or attacking CDPH rather than addressing the substance of the e-cigarette messages (eg, “@ChiPublicHealth You can’t be that stupid. Typical Chicago corruption. Who owns you?”).

Retweeting was highest in the safety category for all tweeters comprising the 86% of the pro-policy tweeting on safety (12 of 14 tweets), and 76% of the anti-policy safety tweeting (271 of 358 tweets). Science was the second most retweeted topic for all tweeters, with retweeting comprising 84% of pro-policy (16 of 19 tweets) and 71.6% of anti-policy (156 of 218 tweets) tweeting. Table 2 shows the number of tweets and retweets by sentiment and category.

Table 2. Themes in e-cigarette tweets mentioning the Chicago Department of Public Health in January, 2014^a.

Theme	Sentiment	Definition	Example tweet	Alln=683 n (%)	Tweet- sn=255	Retweet- sn=428
Safety						
	Pro-policy	e-cigarettes are harmful, foster nicotine addiction, promote smoking	RT @ChiPublicHealth: Electronic cigs contain a dangerous, addictive drug & should be regulated like other nicotine products #ecigtruths htt...	14 (2.0%)	2 (0.8%)	12 (2.8%)
	Anti-policy	e-cigarettes are safer than alternative, promote cessation	@ChiPublicHealth it's not about being safe, it's about being SAFER than the alternative #EcigsSaveLives it's about HARM REDUCTION #Casaa	358 (52.4%)	87 (34.1%)	271 (63.3%)
Lies/propaganda						
	Pro-policy	Propaganda/ lie spread by e-cigarette industry or supporter	N/A	0 (0%)	0 (0%)	0 (0%)
	Anti-policy	Propaganda/ lie spread by health department or other government	@ChiPublicHealth Baseless nonfactual propaganda anyone? So much for public health. #GetAClue #ecigtruths #LiesToldOnTwitter	224 (32.8%)	108 (42.4%)	116 (27.1%)
Science						
	Pro-policy	Studies find some ingredients are carcinogenic, increased use by kids; need more research	@AmerAcadPeds @ChiPublicHealth Time for local & FDA action to protect youth from e-cigs & toxins in both vapor and smoke #SGR50 #putkids1st	19 (2.8%)	3 (1.2%)	16 (3.7%)
	Anti-policy	Science shows e-cigarettes contain only nicotine and water, no dangerous secondhand vapor	@ChiPublicHealth "Vaping: it's not smoking [URL]" No smoke. No carcinogens. No shame. #EcigsSaveLives #Casaa #IMPROOF	218 (31.9%)	62 (24.3%)	156 (36.4%)
Flavor						
	Pro-policy	Sweet flavors are for kids	RT @ChiPublicHealth: "9 Terribly Disturbing Things About Electronic Cigarettes" [URL] via @HuffPostBiz #ecigtruths	2 (0.3%)	0 (0%)	2 (0.5%)
	Anti-policy	Adults like flavors too	@choucair @ChiPublicHealth no one advocates children smoking. that said, my favorite flavor is of strawberries and watermelon. i cant enjoy?	25 (3.7%)	20 (7.8%)	5 (1.2%)
Regulation						
	Pro-policy	Ingredients, look, and use are like cigarettes, should be regulated like cigarettes	RT @IllinoisAFP: @ChiPublicHealth If it looks like a cigarette & contains nicotine like a cigarette, it should be regulated like a cigarett...	44 (6.4%)	10 (3.9%)	34 (7.9%)
	Anti-policy	Regulation is a slippery slope, do not need nanny state	@IllinoisAFP @ChiPublicHealth it looks like a gun, has a trigger like a gun. Let's ban nerf and water guns!	169 (24.7%)	61 (23.9%)	108 (25.2%)
Issue salience						
	Pro-policy	E-cigarettes are an important threat to public health	N/A	0 (0%)	0 (0%)	0 (0%)
	Anti-policy	Health department should focus on more serious health threats	@ChicagosMayor @ChiPublicHealth Y not focus on more free hlth care N staff at the safe passage. Protect kids by actually protecting them.	16 (2.3%)	8 (3.1%)	8 (1.9%)

^aSee [Multimedia Appendix 2](#) for table with URLs included.

Retweet Network

Of the 683 tweets, 62.7% (n=429) were retweets. The majority of the anti-policy tweets were sent January 9th, while anti-policy retweets were frequent on the 9th and again on January 13th, two days before the city council vote. Pro-policy tweets and retweets were infrequent throughout the week. The distribution of anti- and pro-policy tweets and retweets is shown in Figure 1, which also includes non-CDPH social media events related to the Twitter bombing. For example, on January 10th, the Chicago City Council Joint Committee on Finance and Committee on Health and Environmental Protection posted the agenda for their January 13th meeting; the agenda included the e-cigarette policy. Soon after the agenda was posted, Consumer Advocates for Smoke-free Alternatives Association (CASAA), posted an “URGENT CALL TO ACTION” listing strategies to oppose the e-cigarette policy. Figure 1 shows CASAA involvement in a number of the anti-policy social media activities throughout the week.

The retweet network consisted of 259 Twitter users connected to at least one other network member by a retweet link and 47 isolated Twitter users not involved in retweeting. There were 361 asymmetric ties and five mutual ties in the network. Four of the five mutual ties were between the Twitter user labeled “@A” and other network members, and the fifth mutual tie was between @B and @C.

Figure 2 shows the network with nodes sizes representing outdegree centrality; the larger nodes were retweeted more often. Six network members had extremely high outdegree centrality. Specifically, Twitter users represented by nodes @A-@E and @ChiPublicHealth were retweeted 22-59 times; no other network member was retweeted more than 10 times. Four of the central Twitter users were affiliated with formal e-cigarette businesses or advocacy groups.

An examination of the six highly retweeted network members found a single original tweet from @B was retweeted 53 times (RT-1), each time using a different shortened URL, with all URLs directed to the same place. The same strategy was used by @C, whose original tweet (RT-2) was retweeted 24 times using different shortened URLs directed to the same place. Changing the shortened URL in a repeated tweet is a known strategy used to avoid spam detection on Twitter [53]. A total of 77 retweets of tweets from @B and @C were sent using the same strategy; @B and @C were also connected by one of very few strong mutual ties in the network.

*The #publichealth gain is going to be very, very large-biggest likely this century [URL]
@ChiPublicHealth #EcigTruths [RT-1]*

This tweet (RT-1) linked to a newspaper article outlining the debate over e-cigarettes focusing on safety concerns, harm reduction arguments, discussion of laws banning their use, and their increasing popularity (sent by @B; retweeted 53 times).

*@ChiPublicHealth #Please review all of the science before condemning smokers to #QuitOrDie [URL]
#ECigTruths [RT-2]*

This tweet (RT-2) linked to a peer-reviewed journal article about the chemistry of contaminants in e-cigarettes [11] (sent by @C; retweeted 24 times). RT-2 included a link to a peer-reviewed journal article recommended in the CASAA January 10th call to action; the work in the paper was funded by CASAA [11] (see Acknowledgements section of that paper). Overall, 59 retweets, all in the science topic category, linked to this article. The Twitter profile of @B included a link to CASAA, while the profile for @C listed e-cigarette advocacy, but did not include a group affiliation. The Twitter profiles of @B and @C indicated they were not from Chicago. The Twitter profile of the most frequently retweeted network member, @D, lists an e-cigarette retailer as employer; @D sent four original tweets (RT-3 through RT-6) garnering 69 retweets from 59 unique users.

*#ecigs Save lives. NO SMOKE WITHOUT FIRE
@ChiPublicHealth [RT-3]*

RT-3 linked to a photo showing a combustible cigarette on the left and an e-cigarette on the right with a list in the middle comparing the chemicals in smoke and vapor (retweeted 6 times).

Bought this lollipop myself at pharmacy at pediatrician office. #ecigs save lives. Lollipops DO NOT @ChiPublicHealth [RT-4]

RT-4 linked to a photo displaying a nicotine lollipop with a caption comparing the harm reduction effectiveness and reception of nicotine lollipops to e-cigarettes (retweeted 43 times).

[URL] @ChiPublicHealth #ecigs save lives [RT-5]

RT-5 linked to a photo showing a cartoon character saying, “Are you saying that you lie to people about ecigs so that they will not use them even though they work and are safer?” (retweeted 20 times).

[URL] @ChiPublicHealth [RT-6]

RT-6 linked to a photo of a crumpled combustible cigarette with a caption stating e-cigarettes are the best harm reduction alternative to smoking but big corporations are paying the FDA to over-regulate them (not retweeted).

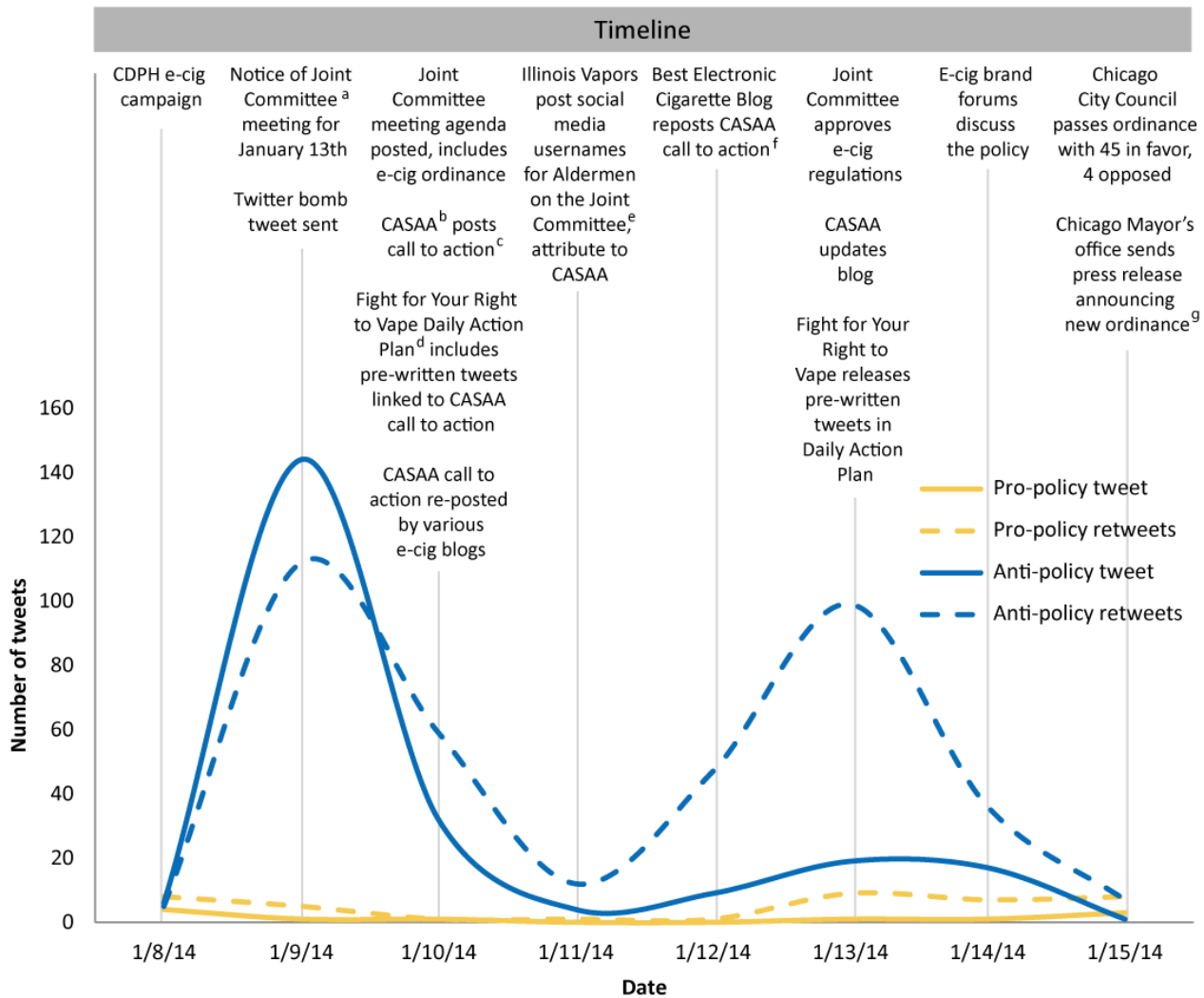
A total of 11 different tweets from @A were retweeted between one and 12 times. The original Twitter bomb tweet, sent by @A, was retweeted nine times. Another central user, @E, is a Twitter account representing an online community of “e-cigarette users and campaigners”. @E sent a single tweet that was retweeted 23 times by 23 different Twitter users. The URL in the tweet was to the published study [11] recommended in the January 10th CASAA call to action; the URL was truncated in all 23 retweets.

RT @E: @ChiPublicHealth New paper shows no danger fm "second hand vape": Great news, you can stop worrying now. [truncated URL]... [RT-7]

The original link pointed to the *BMC Public Health* journal (see RT-2).

Finally, @ChiPublicHealth was also central in the retweet network with 29 pro-policy, and three anti-policy, retweets.

Figure 1. Timeline showing social media activities and distribution of tweets and retweets directed to @ChiPublicHealth in support of and opposing regulation of e-cigarettes in Chicago.



^a Chicago City Council Joint Committee on Finance and Committee on Health and Environmental Protection

^b Consumer Advocates for Smoke-free Alternatives Association

^c CASAA call to action: <http://blog.casaa.org/2014/01/urgent-cta-chicago-e-cigarette-usage.html>

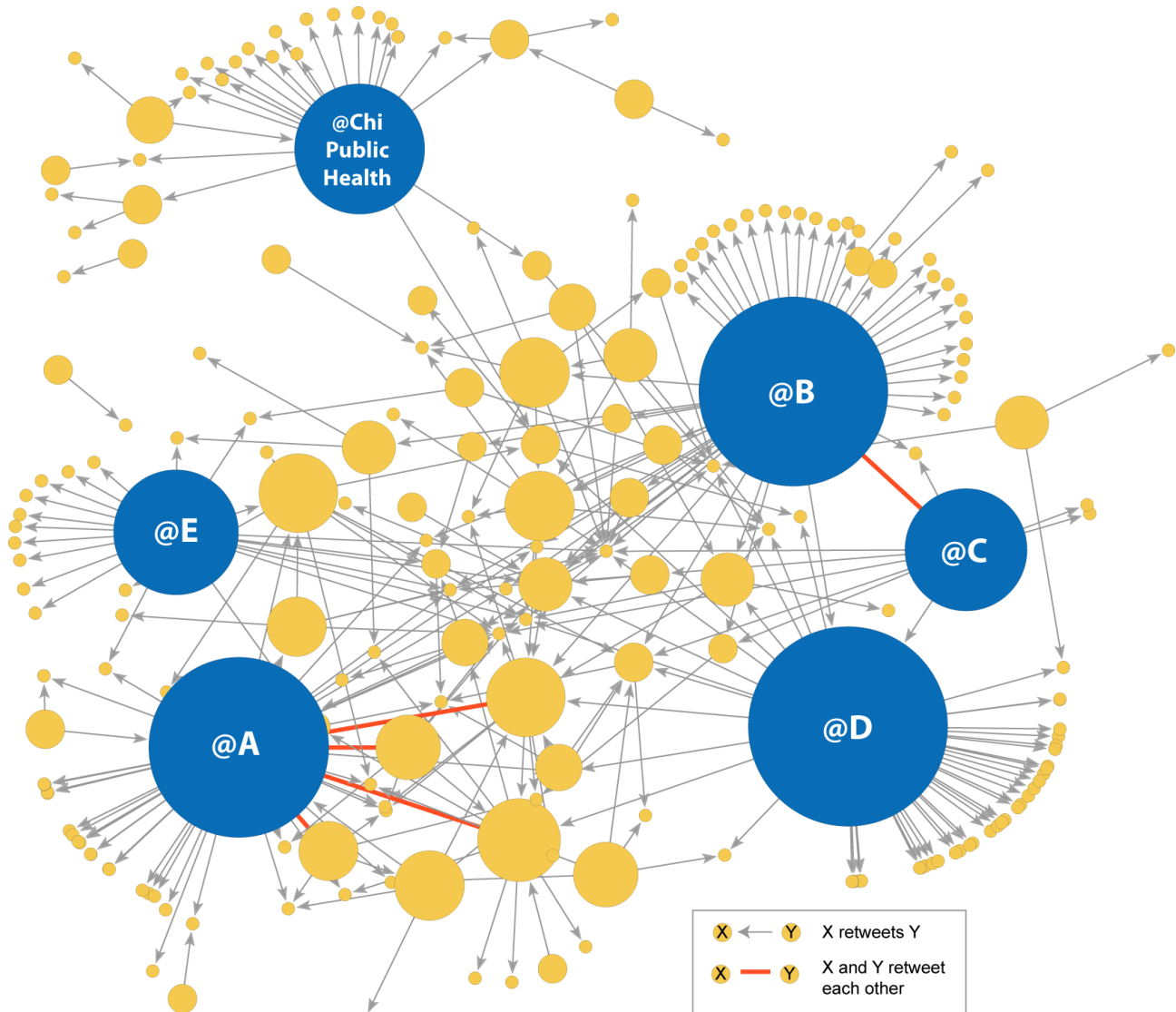
^d Fight for Your Right to Vape Action Plan: <https://www.facebook.com/VapingDailyActionPlan/posts/287318261421539>

^e Illinois Vapors January 11 posts: <https://illinoisvapors.wordpress.com/2014/01/11/>

^f Best Electronic Cigarette Blog post: <http://bestelectroniccigaretteblog.com/casaa-declares-urgent-call-action-chicago>

^g Legislation record #SO2013-6160: <https://chicago.legistar.com/Legislation.aspx>

Figure 2. Retweet network with nodes sized by outdegree, or number of times the Twitter user was retweeted during the week. The 6 most highly retweeted network members shown in blue.



Discussion

Principal Findings

In the week leading up to a vote on local e-cigarette policy by the Chicago City Council, 683 tweets and retweets about e-cigarettes including a mention of @ChiPublicHealth were sent by 307 Twitter users. The majority of the tweets and retweets were against the policy (89.2%, 609/683), and a majority of Twitter users reporting a location in their profile were from states outside Illinois (n=83) or outside the United States (n=48). Twitter users located in Chicago (n=35) were significantly more likely than expected to tweet in favor of the policy.

The CDPH and five Twitter users tweeting against the policy were central in the network of retweeting during this week. Retweet patterns were consistent with past research, which has identified argument quality and source as factors associated with retweeting [58]. In this case, four of the five central retweet network members were affiliated with e-cigarette businesses or advocacy groups, which may have credibility among supporters

of e-cigarette use. Safety and science tweets were retweeted more than lies/propaganda, perhaps due to the lower quality arguments (eg, name calling) in the lies/propaganda tweets.

The majority of tweets appeared to be from legitimate Twitter users who oppose the regulation of e-cigarettes, and at least one advocacy group (CASAA) aiming to organize policy opposition messaging. However, our results suggested that 96 of the 683 tweets (14.1%) sent by 73 of the 307 Twitter users (23.8%) were using an account or tweet strategy consistent with astroturfing. The structure of the retweeting network was also consistent with findings from a study of astroturfing where a small number of accounts were responsible for a large proportion of retweets contributing to trending topics [55].

Recent studies of grassroots efforts found email and phone contact with legislators to have a substantial influence on legislative voting behavior [59,60]. Although there are no studies of social media influence on legislative decision making that we know of, a 2011 study found astroturfing has been successful in influencing public opinion [61]. In addition, a 2011 survey of legislative staffers indicated that Twitter is effective in

reaching legislators [62], and a 2014 study found that constituents lobby via Twitter [63]. During the Twitter bombing of the CDPH, Chicago City Council members were also a target of the Twitter bombing; City Council members Twitter handles were posted January 11th on the Illinois Vapors blog (see Figure 1). Local health departments and local policy-makers in New York City and Los Angeles have also been targeted via social media by e-cigarette advocates. In Chicago, the Twitter anti-policy campaign did not appear to influence the vote, which was 45-4 in favor of regulating e-cigarettes. However, given the widespread use of Twitter by policy-makers and the potential of grassroots efforts to influence legislative decision making, it is increasingly important to understand and address social media policy advocacy strategies and how they may influence development and support of public health policy.

Advocacy efforts that appear to include astroturfing should be fairly easy to detect. To identify accounts associated with astroturfing, review tweets for (1) different shortened URLs pointing to the same place, and (2) Twitter users who have more than one of the characteristics associated with astroturfing (ie, new account, letters and numbers in the username, and few followers). Removing central nodes from a network is a strategy that has been used to disrupt crime and disease transmission networks [64,65], however, to our knowledge it has not yet been studied as a way to address astroturfing. There are two ways to report Twitter users who seem to be involved in astroturfing or spamming, which will prompt Twitter to review the account and possibly remove it: (1) click “report as spam” link appearing with the offending tweet, or (2) post a tweet that includes “@spam @username” where the username of the spammer is included.

The emergence of social media efforts by the tobacco industry [21] and other industries with products promoting risky health behaviors [66] suggest that new social media strategies are needed to combat novel marketing efforts and to increase the presence of public health on social media platforms. In the case

of the Chicago local e-cigarette policy, the number of tweets against the policy was more than 10 times higher than the number of pro-policy tweets despite high policy support from local tweeters. As in offline tobacco use prevention and control, online tobacco use prevention and control may wish to take note of the strategies employed by pro-tobacco interests and adapt them to develop effective counter-marketing. For example, while astroturfing is not an appropriate strategy for public health, tactics employed in astroturfing such as coordinated, widespread, and constant use of specific hashtags or messages (like those supplied by CASAA) to elevate a public health topic on Twitter might work to engage the public around a public health topic. Developing messages that include high quality arguments and originate with reputable sources may aid in increasing message spread [58].

Conclusions

New media marketing strategies present both challenges and opportunities for public health [67]. There is evidence that misinformation spreads easily on the Internet, especially in social media [68,69]. In addition, although evidence of behavior change resulting from social media use is limited and mixed in public health [70,71], emerging successes with public health interventions that engage participants [71,72], and evidence of an association between Facebook content and smoking and alcohol use in adolescents [73], reinforce the importance of understanding social media engagement with substance promoting and health promotion messages. Efforts to develop the evidence base for social media use in public health are underway through funding opportunities such as the National Institutes of Health 2014 Request for Applications, entitled *Using social media to understand and address substance use and addiction* (RFA-CA-14-008 and RFA-CA-14-009). As social media use continues to grow, additional research is needed to better understand how to develop and implement effective pro-health social media campaigns that engage the public to improve health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

E-cigarette tweets sent by the Chicago Department of Public Health (@ChiPublicHealth) on January 8, 2014.

[PDF File (Adobe PDF File), 37KB - [jmir_v16i10e238_app1.pdf](#)]

Multimedia Appendix 2

Themes in e-cigarette tweets mentioning the Chicago Department of Public Health in January, 2014.

[PDF File (Adobe PDF File), 43KB - [jmir_v16i10e238_app2.pdf](#)]

References

1. King BA, Alam S, Promoff G, Arrazola R, Dube SR. Awareness and ever-use of electronic cigarettes among U.S. adults, 2010-2011. *Nicotine Tob Res* 2013 Sep;15(9):1623-1627. [doi: [10.1093/ntr/ntt013](#)] [Medline: [23449421](#)]
2. Centers for Disease Control/Prevention (CDC). Notes from the field: electronic cigarette use among middle and high school students - United States, 2011-2012. *MMWR Morb Mortal Wkly Rep* 2013 Sep 6;62(35):729-730 [FREE Full text] [Medline: [24005229](#)]

3. Choi K, Forster J. Characteristics associated with awareness, perceptions, and use of electronic nicotine delivery systems among young US Midwestern adults. *Am J Public Health* 2013 Mar;103(3):556-561 [FREE Full text] [doi: [10.2105/AJPH.2012.300947](https://doi.org/10.2105/AJPH.2012.300947)] [Medline: [23327246](https://pubmed.ncbi.nlm.nih.gov/23327246/)]
4. Pearson JL, Richardson A, Niaura RS, Vallone DM, Abrams DB. e-Cigarette awareness, use, and harm perceptions in US adults. *Am J Public Health* 2012 Sep;102(9):1758-1766 [FREE Full text] [doi: [10.2105/AJPH.2011.300526](https://doi.org/10.2105/AJPH.2011.300526)] [Medline: [22813087](https://pubmed.ncbi.nlm.nih.gov/22813087/)]
5. Wagener TL, Meier E, Hale JJ, Oliver ER, Warner ML, Driskill LM, et al. Pilot investigation of changes in readiness and confidence to quit smoking after E-cigarette experimentation and 1 week of use. *Nicotine Tob Res* 2014 Jan;16(1):108-114. [doi: [10.1093/ntr/ntt138](https://doi.org/10.1093/ntr/ntt138)] [Medline: [24154511](https://pubmed.ncbi.nlm.nih.gov/24154511/)]
6. Caponnetto P, Campagna D, Cibella F, Morjaria JB, Caruso M, Russo C, et al. Efficiency and Safety of an eElectronic cigAreTte (ECLAT) as tobacco cigarettes substitute: a prospective 12-month randomized control design study. *PLoS One* 2013;8(6):e66317 [FREE Full text] [doi: [10.1371/journal.pone.0066317](https://doi.org/10.1371/journal.pone.0066317)] [Medline: [23826093](https://pubmed.ncbi.nlm.nih.gov/23826093/)]
7. Goniewicz ML, Lingas EO, Hajek P. Patterns of electronic cigarette use and user beliefs about their safety and benefits: an internet survey. *Drug Alcohol Rev* 2013 Mar;32(2):133-140 [FREE Full text] [doi: [10.1111/j.1465-3362.2012.00512.x](https://doi.org/10.1111/j.1465-3362.2012.00512.x)] [Medline: [22994631](https://pubmed.ncbi.nlm.nih.gov/22994631/)]
8. Popova L, Ling PM. Alternative tobacco product use and smoking cessation: a national study. *Am J Public Health* 2013 May;103(5):923-930 [FREE Full text] [doi: [10.2105/AJPH.2012.301070](https://doi.org/10.2105/AJPH.2012.301070)] [Medline: [23488521](https://pubmed.ncbi.nlm.nih.gov/23488521/)]
9. Food & Drug Administration. News & Events. 2014. Electronic cigarettes (e-cigarettes) URL: <http://www.fda.gov/newsevents/publichealthfocus/ucm172906.htm> [accessed 2014-06-16] [WebCite Cache ID 6QNH8v8EX]
10. Hua M, Alfi M, Talbot P. Health-related effects reported by electronic cigarette users in online forums. *J Med Internet Res* 2013;15(4):e59 [FREE Full text] [doi: [10.2196/jmir.2324](https://doi.org/10.2196/jmir.2324)] [Medline: [23567935](https://pubmed.ncbi.nlm.nih.gov/23567935/)]
11. Burstyn I. Peering through the mist: systematic review of what the chemistry of contaminants in electronic cigarettes tells us about health risks. *BMC Public Health* 2014;14:18 [FREE Full text] [doi: [10.1186/1471-2458-14-18](https://doi.org/10.1186/1471-2458-14-18)] [Medline: [24406205](https://pubmed.ncbi.nlm.nih.gov/24406205/)]
12. Cheng T. Chemical evaluation of electronic cigarettes. *Tob Control* 2014 May;23 Suppl 2:ii11-ii17 [FREE Full text] [doi: [10.1136/tobaccocontrol-2013-051482](https://doi.org/10.1136/tobaccocontrol-2013-051482)] [Medline: [24732157](https://pubmed.ncbi.nlm.nih.gov/24732157/)]
13. Orr MS. Electronic cigarettes in the USA: a summary of available toxicology data and suggestions for the future. *Tob Control* 2014 May;23 Suppl 2:ii18-ii22 [FREE Full text] [doi: [10.1136/tobaccocontrol-2013-051474](https://doi.org/10.1136/tobaccocontrol-2013-051474)] [Medline: [24732158](https://pubmed.ncbi.nlm.nih.gov/24732158/)]
14. Callahan-Lyon P. Electronic cigarettes: human health effects. *Tob Control* 2014 May;23 Suppl 2:ii36-ii40 [FREE Full text] [doi: [10.1136/tobaccocontrol-2013-051470](https://doi.org/10.1136/tobaccocontrol-2013-051470)] [Medline: [24732161](https://pubmed.ncbi.nlm.nih.gov/24732161/)]
15. Kosmider L, Sobczak A, Fik M, Knysak J, Zaciera M, Kurek J, et al. Carbonyl compounds in electronic cigarette vapors: effects of nicotine solvent and battery output voltage. *Nicotine Tob Res* 2014 Oct;16(10):1319-1326. [doi: [10.1093/ntr/ntu078](https://doi.org/10.1093/ntr/ntu078)] [Medline: [24832759](https://pubmed.ncbi.nlm.nih.gov/24832759/)]
16. Dutra LM, Glantz SA. Electronic cigarettes and conventional cigarette use among U.S. adolescents: a cross-sectional study. *JAMA Pediatr* 2014 Jul;168(7):610-617 [FREE Full text] [doi: [10.1001/jamapediatrics.2013.5488](https://doi.org/10.1001/jamapediatrics.2013.5488)] [Medline: [24604023](https://pubmed.ncbi.nlm.nih.gov/24604023/)]
17. Kim AE, Arnold KY, Makarenko O. E-cigarette advertising expenditures in the U.S., 2011-2012. *Am J Prev Med* 2014 Apr;46(4):409-412. [doi: [10.1016/j.amepre.2013.11.003](https://doi.org/10.1016/j.amepre.2013.11.003)] [Medline: [24650844](https://pubmed.ncbi.nlm.nih.gov/24650844/)]
18. Duke JC, Lee YO, Kim AE, Watson KA, Arnold KY, Nonnemaker JM, et al. Exposure to electronic cigarette television advertisements among youth and young adults. *Pediatrics* 2014 Jul;134(1):e29-e36. [doi: [10.1542/peds.2014-0269](https://doi.org/10.1542/peds.2014-0269)] [Medline: [24918224](https://pubmed.ncbi.nlm.nih.gov/24918224/)]
19. Freeman B. New media and tobacco control. *Tob Control* 2012 Mar;21(2):139-144. [doi: [10.1136/tobaccocontrol-2011-050193](https://doi.org/10.1136/tobaccocontrol-2011-050193)] [Medline: [22345236](https://pubmed.ncbi.nlm.nih.gov/22345236/)]
20. Ayers JW, Ribisl KM, Brownstein JS. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. *Am J Prev Med* 2011 Apr;40(4):448-453. [doi: [10.1016/j.amepre.2010.12.007](https://doi.org/10.1016/j.amepre.2010.12.007)] [Medline: [21406279](https://pubmed.ncbi.nlm.nih.gov/21406279/)]
21. De AM, Hastings G, Angus K, Dixon D, Purves R. Cancer Research UK. 2013. The marketing of electronic cigarettes in the UK URL: http://www.cancerresearchuk.org/prod_consump/groups/cr_common/@nre/@pol/documents/generalcontent/cr_115991.pdf [accessed 2014-06-15] [WebCite Cache ID 6QNHPU5DD]
22. Brown JE, Luo W, Isabelle LM, Pankow JF. Candy flavorings in tobacco. *N Engl J Med* 2014 Jun 5;370(23):2250-2252. [doi: [10.1056/NEJMc1403015](https://doi.org/10.1056/NEJMc1403015)] [Medline: [24805984](https://pubmed.ncbi.nlm.nih.gov/24805984/)]
23. Durmowicz EL. The impact of electronic cigarettes on the paediatric population. *Tob Control* 2014 May;23 Suppl 2:ii41-ii46 [FREE Full text] [doi: [10.1136/tobaccocontrol-2013-051468](https://doi.org/10.1136/tobaccocontrol-2013-051468)] [Medline: [24732163](https://pubmed.ncbi.nlm.nih.gov/24732163/)]
24. Food & Drug Administration. 2014. Tobacco products: extending authorities to additional tobacco products URL: <http://www.fda.gov/TobaccoProducts/Labeling/ucm388395.htm> [accessed 2014-06-16] [WebCite Cache ID 6QNHBrCb3]
25. American Nonsmokers' Rights Foundation. 2014. US state and local laws regulating use of electronic cigarettes URL: <http://www.no-smoke.org/pdf/ecigslaws.pdf> [accessed 2014-06-16] [WebCite Cache ID 6QNHvukl]
26. City of Chicago. Mayor's Press Office. 2014. City council approves ordinance to regulate e-cigarettes as tobacco products URL: <http://www.cityofchicago.org/city/en/depts/cdph/provdrs/boh/news/2014/jan/city-council-approves-ordinance-to-regulate-e-cigarettes-as-toba.html> [accessed 2014-06-16] [WebCite Cache ID 6QNHnN7VN]

27. Office of the City Clerk. City of Chicago. 2014. Legislation URL: <https://chicago.legistar.com/Legislation.aspx> [accessed 2014-06-15] [WebCite Cache ID 6QM28nWII]
28. Centers for Disease Control and Prevention. National Public Health Performance Standards Program (NPHPSP). 2010. 10 Essential Public Health Services URL: <http://www.cdc.gov/nphpsp/essentialservices.html> [accessed 2013-09-23] [WebCite Cache ID 6JqfMLG6D]
29. Public Health Accreditation Board. 2011. Public Health Accreditation Board standards: an overview URL: <http://www.phaboard.org/wp-content/uploads/PHAB-Standards-and-Measures-Version-1.0.pdf> [accessed 2013-11-16] [WebCite Cache ID 6LAqTGqLC]
30. Mays GP, McHugh MC, Shim K, Lenaway D, Halverson PK, Moonesinghe R, et al. Getting what you pay for: public health spending and the performance of essential public health services. *J Public Health Manag Pract* 2004 Oct;10(5):435-443. [Medline: 15552769]
31. Harris JK, Mueller NL, Snider D, Haire-Joshu D. Local health department use of Twitter to disseminate diabetes information, United States. *Prev Chronic Dis* 2013;10:E70 [FREE Full text] [doi: 10.5888/pcd10.120215] [Medline: 23639765]
32. Harris JK, Mueller NL, Snider D. Social media adoption in local health departments nationwide. *Am J Public Health* 2013 Sep;103(9):1700-1707. [doi: 10.2105/AJPH.2012.301166] [Medline: 23865660]
33. Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. *BMC Public Health* 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]
34. Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. Use of Twitter among local health departments: an analysis of information sharing, engagement, and action. *J Med Internet Res* 2013;15(8):e177 [FREE Full text] [doi: 10.2196/jmir.2775] [Medline: 23958635]
35. Heldman AB, Schindelar J. Social media engagement and public health communication: implications for public health organizations being truly "social". *Public Health Reviews* 2013;35(1):1-1 [FREE Full text]
36. Ratkiewicz J, Conover M, Meiss M, Goncalves B, Flammini A, Menczer F. ICWSM. 2011. Detecting and tracking political abuse in social media URL: http://truthy.indiana.edu/site_media/pdfs/ratkiewicz_icwsm2011_truthy.pdf [accessed 2014-06-16] [WebCite Cache ID 6QNIMMGNS]
37. Fallin A, Grana R, Glantz SA. 'To quarterback behind the scenes, third-party efforts': the tobacco industry and the Tea Party. *Tob Control* 2014 Jul;23(4):322-331 [FREE Full text] [doi: 10.1136/tobaccocontrol-2012-050815] [Medline: 23396417]
38. Lee CW. The roots of astroturfing. *Contexts* 2010 Feb;9(1):73-75 [FREE Full text] [doi: 10.1525/ctx.2010.9.1.73]
39. Lyon TP, Maxwell JW. Astroturf: interest group lobbying and corporate strategy. *J Economics Management Strategy* 2004 Dec;13(4):561-597. [doi: 10.1111/j.1430-9134.2004.00023.x]
40. The Center for Media and Democracy. 2014. Sourcewatch: Astroturf URL: <http://www.sourcewatch.org/index.php/Astroturf> [accessed 2014-06-16] [WebCite Cache ID 6QNJhG3NO]
41. Samuels B, Glantz SA. The politics of local tobacco control. *JAMA* 1991 Oct 16;266(15):2110-2117. [Medline: 1920699]
42. Mustafaraj E, Metaxas PT. From obscurity to prominence in minutes: political speech and real-time search. In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line. Raleigh, NC; 2010 Presented at: Web Science Conference; April 26-27, 2010; Raleigh, NC, USA URL: http://journal.webscience.org/317/2/websci10_submission_89.pdf
43. Lo CY. Astroturf versus grass roots: scenes from early Tea Party mobilization. In: Rosenthal L, Trost C, editors. *Steep: the Precipitous Rise of the Tea Party*. Berkeley, CA: University of California Press; 2012:98-129.
44. Bruns A, Burgess JE. The use of Twitter hashtags in the formation of ad hoc publics. 2011 Presented at: 6th European Consortium for Political Research General Conference; August, 2011; University of Iceland, Reykjavik URL: <http://eprints.qut.edu.au/46515/>
45. Suh B, Hong L, Pirolli P, Chi EH. Want to be retweeted? large scale analytics on factors impacting retweet in Twitter networks. 2010 Aug 20 Presented at: IEEE Second International Conference on Social Computing (SocialCom); 2010; Minneapolis, MN, USA p. 177-184.
46. Naveed N, Gottron T, Kunegis J, Alhadi AC. Bad news travel fast: a content-based analysis of interestingness on Twitter. 2011 Jun 14 Presented at: ACM Web Science; 2011; Koblenz, Germany p. 2011.
47. Twitter. What are @replies and mentions?. 2014. URL: <https://support.twitter.com/articles/14023-what-are-replies-and-mentions> [accessed 2014-06-15] [WebCite Cache ID 6QM217D36]
48. Gentry J. The Comprehensive R Archive Network. 2013. twitter: R based Twitter client, v. 1.1.7 URL: <http://cran.r-project.org/web/packages/twitter/index.html> [accessed 2014-09-29] [WebCite Cache ID 6SxZkbMW8]
49. Hansen D, Shneiderman B, Smith MA. Analyzing social media networks with NodeXL: insights from a connected world. Burlington, MA: Morgan Kaufmann; Sep 10, 2010.
50. Smith M, Milic-Frayling N, Shneiderman B, Mendes Rodrigues E, Leskovec J, Dunne C. NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010.: Social Media Research Foundation URL: <http://www.smrfoundation.org/> [accessed 2014-09-29] [WebCite Cache ID 6SxkKKWD7]
51. Mrvar A, Batagelj V. Pajek64. URL: <http://pajek.imfm.si/doku.php> [accessed 2014-09-29] [WebCite Cache ID 6SxkQ3fLe]
52. De Nooy W, Mrvar A, Batagelj V. Exploratory social network analysis with Pajek. Exploratory social network analysis with Pajek: Cambridge University Press; 2011.

53. Wang AH. Don't follow me: spam detection in Twitter. : IEEE; 2010 Presented at: International Conference on Security and Cryptography (SECRYPT); 2010; Athens, Greece.
54. Yardi S, Romero D, Schoenebeck G, Boyd D. Detecting spam in a Twitter network. *First Monday* 2009 Dec 30;15(1):1-1. [doi: [10.5210/fm.v15i1.2793](https://doi.org/10.5210/fm.v15i1.2793)]
55. Yu LL, Asur S, Huberman BA. Artificial inflation: the real story of trends and trend-setters in Sina Weibo. : IEEE; 2012 Presented at: International Conference on Privacy, Security, Risk and Trust (PASSAT) and International Conference on Social Computing (SocialCom); 2012; Amsterdam, Netherlands.
56. Luke DA. Viewing dissemination through a network lens. In: Brownson RC, editor. *Dissemination and Implementation Research in Health: Translating Science to Practice*. USA: Oxford University Press; 2012.
57. Harris JK. Communication ties across the national network of local health departments. *Am J Prev Med* 2013 Mar;44(3):247-253. [doi: [10.1016/j.amepre.2012.10.028](https://doi.org/10.1016/j.amepre.2012.10.028)] [Medline: [23415121](https://pubmed.ncbi.nlm.nih.gov/23415121/)]
58. Ha S, Ahn J. Why are you sharing others' Tweets?: the impact of argument quality and source credibility on information sharing behavior. 2011 Presented at: International Conference on Information Systems; 2011; Shanghai URL: <http://aisel.aisnet.org/icis2011/proceedings/humanbehavior/4/>
59. Bergan DE, Cole RT. Call Your Legislator: a field experimental study of the impact of a constituency mobilization campaign on legislative voting. *Polit Behav* 2014 Jun 1:1. [doi: [10.1007/s11109-014-9277-1](https://doi.org/10.1007/s11109-014-9277-1)]
60. Henriksen C, Hansen S, Nordgaard-Lassen I, Rikardt Anderson J, Madsen P. Erratum for PMID 21180585. *Therap Adv Gastroenterol* 2012 Sep;5(5):371 [FREE Full text] [doi: [10.1177/1756283X10363751](https://doi.org/10.1177/1756283X10363751)] [Medline: [22973420](https://pubmed.ncbi.nlm.nih.gov/22973420/)]
61. Cho CH, Martens ML, Kim H, Rodrigue M. Astroturfing global warming: it isn't always greener on the other side of the fence. *J Bus Ethics* 2011 Jul 3;104(4):571-587. [doi: [10.1007/s10551-011-0950-6](https://doi.org/10.1007/s10551-011-0950-6)]
62. Bush M. Edelman. 2011. Social media's ability to influence policymakers grows globally URL: <http://www.edelman.com/news/social-medias-ability-to-influence-policymakers-grows-globally/> [accessed 2014-06-16] [WebCite Cache ID [6QNLoXvgv](https://www.webcitation.org/6QNLoXvgv)]
63. Hemphill L, Roback AJ. Tweet acts: how constituents lobby congress via Twitter. 2014 Presented at: 17th ACM Conference on Computer Supported Cooperative Work & Social Computing; 2014; Baltimore, MD, USA p. 1200-1210 URL: <http://dl.acm.org/citation.cfm?id=2531735>
64. Valente TW. Network interventions. *Science* 2012 Jul 6;337(6090):49-53. [doi: [10.1126/science.1217330](https://doi.org/10.1126/science.1217330)] [Medline: [22767921](https://pubmed.ncbi.nlm.nih.gov/22767921/)]
65. Morselli C. *Crime and networks*. New York, NY: Routledge; 2013.
66. Nicholls J. Everyday, everywhere: alcohol marketing and social media--current trends. *Alcohol Alcohol* 2012;47(4):486-493 [FREE Full text] [doi: [10.1093/alcalc/ags043](https://doi.org/10.1093/alcalc/ags043)] [Medline: [22532575](https://pubmed.ncbi.nlm.nih.gov/22532575/)]
67. Ribisl KM, Jo C. Tobacco control is losing ground in the Web 2.0 era: invited commentary. *Tob Control* 2012 Mar;21(2):145-146. [doi: [10.1136/tobaccocontrol-2011-050360](https://doi.org/10.1136/tobaccocontrol-2011-050360)] [Medline: [22345237](https://pubmed.ncbi.nlm.nih.gov/22345237/)]
68. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
69. Sudau F, Friede T, Grabowski J, Koschack J, Makedonski P, Himmel W. Sources of information and behavioral patterns in online health forums: observational study. *J Med Internet Res* 2014;16(1):e10 [FREE Full text] [doi: [10.2196/jmir.2875](https://doi.org/10.2196/jmir.2875)] [Medline: [24425598](https://pubmed.ncbi.nlm.nih.gov/24425598/)]
70. Chou WY, Prestin A, Lyons C, Wen KY. Web 2.0 for health promotion: reviewing the current evidence. *Am J Public Health* 2013 Jan;103(1):e9-18. [doi: [10.2105/AJPH.2012.301071](https://doi.org/10.2105/AJPH.2012.301071)] [Medline: [23153164](https://pubmed.ncbi.nlm.nih.gov/23153164/)]
71. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelandotte C. Are health behavior change interventions that use online social networks effective? A systematic review. *J Med Internet Res* 2014;16(2):e40 [FREE Full text] [doi: [10.2196/jmir.2952](https://doi.org/10.2196/jmir.2952)] [Medline: [24550083](https://pubmed.ncbi.nlm.nih.gov/24550083/)]
72. Chiu M, Chang S, Chang Y, Chu HH, Chen CCH, Hsiao FH, et al. Playful bottle: a mobile social persuasion system to motivate healthy water intake. 2009 Presented at: Proceedings of the 11th International Conference on Ubiquitous Computing; 2009; Orlando, FL p. 185-194 URL: http://comp.utm.my/rohayanti/files/2013/09/S02-PlayfulBottle_Motivate-Healthy-Water-Intake1.pdf
73. Huang GC, Unger JB, Soto D, Fujimoto K, Pentz MA, Jordan-Marsh M, et al. Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. *J Adolesc Health* 2014 May;54(5):508-514. [doi: [10.1016/j.jadohealth.2013.07.001](https://doi.org/10.1016/j.jadohealth.2013.07.001)] [Medline: [24012065](https://pubmed.ncbi.nlm.nih.gov/24012065/)]

Abbreviations

CASAA: Consumer Advocates for Smoke-free Alternatives Association

CDPH: Chicago Department of Public Health

FDA: Food and Drug Administration

LHD: local health departments

Edited by G Eysenbach; submitted 18.06.14; peer-reviewed by P Talbot; comments to author 04.09.14; accepted 09.09.14; published 16.10.14.

Please cite as:

Harris JK, Moreland-Russell S, Choucair B, Mansour R, Staub M, Simmons K

Tweeting for and Against Public Health Policy: Response to the Chicago Department of Public Health's Electronic Cigarette Twitter Campaign

J Med Internet Res 2014;16(10):e238

URL: <http://www.jmir.org/2014/10/e238/>

doi: [10.2196/jmir.3622](https://doi.org/10.2196/jmir.3622)

PMID: [25320863](https://pubmed.ncbi.nlm.nih.gov/25320863/)

©Jenine K Harris, Sarah Moreland-Russell, Bechara Choucair, Raed Mansour, Mackenzie Staub, Kendall Simmons. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 16.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Rapid Grading of Fundus Photographs for Diabetic Retinopathy Using Crowdsourcing

Christopher J Brady^{1,2}, MD; Andrea C Villanti³, MPH, PhD; Jennifer L Pearson³, MPH, PhD; Thomas R Kirchner³, PhD; Omesh P Gupta¹, MBA, MD; Chirag P Shah⁴, MPH, MD

¹Wills Eye Hospital, Retina Service: Mid Atlantic Retina, Philadelphia, PA, United States

²Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, MD, United States

³Schroeder Institute for Tobacco Research and Policy Studies, Legacy, Washington, DC, United States

⁴Ophthalmic Consultants of Boston, Boston, MA, United States

Corresponding Author:

Christopher J Brady, MD

Wilmer Eye Institute

Johns Hopkins University School of Medicine

600 N Wolfe St.

Maumenee 711

Baltimore, MD, 21287

United States

Phone: 1 (410) 502 2789

Fax: 1 (443) 287 8343

Email: brady@jhmi.edu

Abstract

Background: Screening for diabetic retinopathy is both effective and cost-effective, but rates of screening compliance remain suboptimal. As screening improves, new methods to deal with screening data may help reduce the human resource needs. Crowdsourcing has been used in many contexts to harness distributed human intelligence for the completion of small tasks including image categorization.

Objective: Our goal was to develop and validate a novel method for fundus photograph grading.

Methods: An interface for fundus photo classification was developed for the Amazon Mechanical Turk crowdsourcing platform. We posted 19 expert-graded images for grading by Turkers, with 10 repetitions per photo for an initial proof-of-concept (Phase I). Turkers were paid US \$0.10 per image. In Phase II, one prototypical image from each of the four grading categories received 500 unique Turker interpretations. Fifty draws of 1-50 Turkers were then used to estimate the variance in accuracy derived from randomly drawn samples of increasing crowd size to determine the minimum number of Turkers needed to produce valid results. In Phase III, the interface was modified to attempt to improve Turker grading.

Results: Across 230 grading instances in the normal versus abnormal arm of Phase I, 187 images (81.3%) were correctly classified by Turkers. Average time to grade each image was 25 seconds, including time to review training images. With the addition of grading categories, time to grade each image increased and percentage of images graded correctly decreased. In Phase II, area under the curve (AUC) of the receiver-operator characteristic (ROC) indicated that sensitivity and specificity were maximized after 7 graders for ratings of normal versus abnormal (AUC=0.98) but was significantly reduced (AUC=0.63) when Turkers were asked to specify the level of severity. With improvements to the interface in Phase III, correctly classified images by the mean Turker grade in four-category grading increased to a maximum of 52.6% (10/19 images) from 26.3% (5/19 images). Throughout all trials, 100% sensitivity for normal versus abnormal was maintained.

Conclusions: With minimal training, the Amazon Mechanical Turk workforce can rapidly and correctly categorize fundus photos of diabetic patients as normal or abnormal, though further refinement of the methodology is needed to improve Turker ratings of the degree of retinopathy. Images were interpreted for a total cost of US \$1.10 per eye. Crowdsourcing may offer a novel and inexpensive means to reduce the skilled grader burden and increase screening for diabetic retinopathy.

(*J Med Internet Res* 2014;16(10):e233) doi:[10.2196/jmir.3807](https://doi.org/10.2196/jmir.3807)

KEYWORDS

diabetic retinopathy; telemedicine; fundus photography; crowdsourcing; Amazon Mechanical Turk

Introduction

Since early diabetic retinopathy (DR) is often asymptomatic, detection of disease at this stage is either incidental or by deliberate screening. Screening for DR is both effective and cost-effective [1-4], but adherence rates to published guidelines for screening for DR are low, with only 35-60% of diabetic patients receiving an annual dilated fundus examination in the United States [5-8]. As a way to increase adherence, telehealth screening using non-mydratic fundus photography and remote interpretation is increasing, especially in rural and remote settings [9-11]. Early diagnosis of DR and institution of appropriate therapy represents an enormous opportunity to prevent vision loss in a young, working-age demographic [3,4]. Telehealth, in particular, may be a way to control provider, payer, and societal costs.

Among the costs of a telehealth program are the fundus camera, the telehealth software package, and the human resources needed for image acquisition and interpretation. Fundus photo interpretation costs in diabetic retinopathy screening may be high given the labor-intensive interpretation protocols and the need to interpret multiple images per patient. Computerized, semi-automated image analysis techniques have been developed that may be able to reduce physician workload and screening costs [12-14]; however, these methods are not FDA-approved, nor in wide use clinically at this time. If telehealth continues to expand, low-cost methods will be needed to interpret the large volume of fundus images expected with rising incidence of diabetes, especially in resource-poor settings and in large public health screenings.

Crowdsourcing is defined by Brabham as “an online, distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organizational goals” [15]. A subset of crowdsourcing, which he terms “distributed-human-intelligence tasking”, can involve subdividing larger tasks into small portions and then recruiting a group of individuals to each complete these small portions, and only collectively, the entire task [15]. The use of crowdsourcing in biomedical research is in its infancy, though some groups have used this method in public health research [16] and to interpret medical imaging. For example, malaria researchers have used a Web-based game to recruit untrained, anonymous volunteers to tag and count malaria parasites on digital images of blood smears [17]. The investigators showed that by combining the analyses of several users, they were able to achieve similar accuracy rates to expert microscopists. Crowdsourcing has recently been used to categorize a number of fundus photos with a variety of diagnoses as normal or abnormal [18]. In a trial conducted in the United Kingdom using untrained graders, the sensitivity was $\geq 96\%$ for normal versus severely abnormal and between 61-79% for normal versus mildly abnormal [18].

The current research uses diabetic retinopathy as the test condition to explore whether a crowdsourcing interface can be

used to train workers to classify human fundus photos as normal or abnormal and subsequently conduct diagnostic grading of images [19]. This project estimates the validity and reliability of crowdsourced grading of images for diabetic retinopathy, compared to the criterion standard of expert grading. Our hypothesis is that crowdsourced grading of fundus photography interpretation can be rapid, accurate, and reliable in the screening for diabetic retinopathy.

Methods

Crowdsourcing


An interface for fundus photo classification was developed for the Amazon Mechanical Turk (AMT [20]) crowdsourcing platform (Figure 1). AMT is an online labor market that allows access to thousands of people who can quickly accomplish small, discrete tasks for small amounts of money. Typical AMT tasks include tagging photos, translating words, or writing very short articles for websites. AMT has also been used to annotate photos of the tobacco point-of-sale retail environment [21], evaluate oral health promotion materials [22], investigate the relationship between depression and stigma [23], assess people’s reactions to frightening anti-smoking campaigns [24], and evaluate public awareness of ovarian cancer [25], among many other research-orientated applications [26-28]. Amazon Mechanical Turk has its own vocabulary used by AMT workers (Turkers) and AMT task administrators (Requestors). A Human Intelligence Task (HIT) is a small job that may be performed in a matter of seconds or minutes and, once the work is approved by the requestor, may pay US \$0.01-\$0.25 or more per task depending on the complexity of the HIT. A group of HITs is called a “batch” and is made up of similar HITs. Depending on the complexity of the task and the payment offered by the Requestor, a batch is often completed within minutes or hours of posting.

AMT is a reputation-based economy such that Turkers may only access the most desirable HITs once they have a sufficient track record of previously accepted work [29]. High quality Turkers may avoid a new Requestor’s HITs until the Requestor has demonstrated his or her own fairness in approving and rejecting work. Indeed, a Turker’s reputation will suffer following rejection of even a small number of HITs. AMT is a complex ecosystem in which both high-quality work on the part of the Turkers and fairness on the part of the Requestor are rewarded.

Turkers perform their work anonymously, but demographic studies have been conducted. In a survey of 1000 Turkers, Ipeirotis found that 46.8% of Turkers are located in the United States, 34% are in India, and the remaining 19.2% were from 64 other countries [30]. The majority of workers in the United States were women, most of whom reported AMT as a source of supplemental income, whereas in the majority of workers in India were men, and reported AMT as their primary source of income. Across nations, Turkers were younger and better educated than the general population [30].

Figure 1. Screenshot of the Amazon Mechanical Turk Web interface for fundus photo grading.

This is a photo of the inside of the eye. We are looking to label eyes as healthy or unhealthy with respect to diabetes. This task is difficult, but all good faith efforts will be accepted. Please answer both questions even if you are not sure. To close pop-out, move cursor left, or [tap here](#) on iPad.



Hover for examples:

[Normal](#)

[Abnormal - Mild](#)

[Abnormal - Moderate](#)

[Abnormal - Severe Example 1](#)

[Abnormal - Severe Example 2](#)

[Abnormal - Severe Example 3](#)

1) Rate this eye:
NOTE: The color of the photo is not a marker of disease.

Normal (Healthy)

Abnormal (mild, moderate, severe disease)

2) Is this photo good enough quality to grade?

Yes

No

Please leave any comments or feedback you have about this HIT or project:

Design

For the current study, the United Kingdom national screening program grading scale [31] was used by 2 attending retinal surgeons (OPG, CPS) to categorize 19 anonymized teaching images. The same AMT interface and resolution of images used by the Turkers was used by the expert graders. Consensus was reached by discussion on images for which there was disagreement. For the purposes of the study, terms from the UK grading scale were translated into plain language: “background” retinopathy was called “mild”, “preproliferative” was called “moderate”, and “proliferative” was called “severe”. “Maculopathy” was defined as abnormal on a training image with otherwise moderate disease but was not coded separately. The AMT interface was designed to provide training on grading of DR within each HIT. This training included 6 images annotated with the salient features of each level of retinopathy in plain language. Turkers were presented with the following text: “This is a photo of the inside of the eye. We are looking to label eyes as healthy or unhealthy with respect to diabetes. Rate this eye.” Turkers could hover their mouse over 6 adjacent training images (1 normal, 1 mild, 1 moderate, 3 severe) while reviewing the active test image. This layout allowed for all of the training and grading to occur in one browser window. More examples of proliferative/severe disease were provided due to the heterogeneity of findings constituting this level of disease. There were no restrictions on the type of device or display/monitor used by Turkers to complete the task.

Phase I

In the first phase of the study, the 19 images were posted to AMT for grading by Turkers, with 10 repetitions per photo for an initial proof-of-concept. Four photos were re-posted with this initial batch for 10 repetitions to assess intragrader reliability. Turkers were paid US \$0.10 per image, and a 10%

commission was paid to Amazon. In order to be eligible to view and complete the HITs, Turkers needed to have successfully completed 100 prior HITs and have an overall HIT approval rate of 97%.

In the initial batch, HITs were posted asking Turkers to grade images as normal (0) versus abnormal (1). In subsequent batches, Turkers were asked to grade the same 19 images using three categories (normal=0, mild to moderate=1, and severe=2) and then four categories (normal=0, mild=1, moderate=2, and severe=3). Percentage of images correctly classified was calculated. Sensitivity and specificity were calculated for all batches, collapsing all abnormal categories together for three- and four-category grading tasks using Stata 13. A worker consensus grade was assigned to each image based on the mode of the 10 Turker scores. Consensus grade using the mean of the 10 Turker scores was also calculated on an exploratory basis. For the two category tasks, a mean score <0.50 was defined as normal, and ≥ 0.50 was abnormal. For the three category tasks, <0.50 was defined as normal, ≥ 0.5 to <1.5 was defined as mild to moderate, and ≥ 1.5 was defined as severe. In the four category tasks, <0.50 was defined as normal, ≥ 0.5 to <1.5 was defined as mild, ≥ 1.5 to <2.5 was defined as moderate, and ≥ 2.5 was defined as severe.

Phase II

The purpose of Phase II was two-fold. First, Phase II sought to evaluate the emergent ability of the crowd to accurately distinguish between different levels of retinopathy, based on the idea that larger numbers of raters would increasingly coalesce around the correct answer. Second, Phase II sought to identify and ultimately confirm the threshold beyond which the contribution of one more rater would cease to provide additional information. To accomplish this, one prototypical image from each of the four grading categories was submitted to undergo

500 unique Turker interpretations. Jackknife resampling methods were then used to draw random subsamples from this “population” of 500 Turkers, beginning with 50 random samples of 2 Turkers, then 50 random samples of 3 Turkers, and so forth [21,32]. This made it possible to estimate the variance in accuracy derived from each randomly drawn sample of raters at each crowd size and to compute area under the curve (AUC) of the receiver-operator characteristic (ROC) curve, indicating their performance relative to the expert grading.

Phase III

In Phase III of the study, three additional iterative batches of the 19 images were run in an effort to improve Turker grading using the four diagnostic categories. The first batch used Turker feedback from all previous batches to modify the training image set. The second batch used more stringent criteria for Turkers in addition to the modified training images. A minimum of 500 completed and approved HITs was required as was an overall HIT approval rate of 99%. The third batch was conducted using Turkers holding an Amazon designation of “Photo Moderation Master” and raised the compensation to US \$0.15 per image, in addition to the changes made for batches 1 and 2. The criteria necessary to achieve this designation are not published, but Masters are “elite groups of Workers who have demonstrated accuracy on specific types of HITs on the Mechanical Turk marketplace. Workers achieve a Masters distinction by consistently completing HITs of a certain type with a high degree of accuracy across a variety of Requesters” [33]. As in Phase I, percent correctly classified, worker consensus score, and average time to complete the HITs were estimated for these three iterative batches.

The Wills Eye Institute Institutional Review Board ruled that approval was not required for this study.

Results

Phase I

Two expert graders (OPG, CPS) coded 12 images as abnormal and 7 as normal (Table 1). Each of the three Phase I batches

consisted of 23 photos (19 unique, 4 duplicates) with 10 unique graders for a total of 230 grading instances. Time-to-complete Turker grading of images varied with the number of grading categories. Two category (normal/abnormal) grading was completed in 20 minutes, three-category grading in 3 hours, and four-category grading in 2 hours. Because the images were interpreted rapidly and workers could complete as many or as few of the HITs as desired, most of the four duplicate images were rated by unique Turkers and therefore, we were unable to assess intragrader reliability.

Across 230 grading instances of unique images in the two-category HITs (normal vs. abnormal) of Phase I, 187 (81.3%) of the images were correctly classified by Turkers (Table 1). Sensitivity and specificity were 93.6% and 67.8% respectively using individual Turker scores. Sensitivity and specificity were 100% and 71.4% respectively using Turker consensus scores. Average time to grade each image was 25 seconds, including time to review training images. At US \$0.11 per grading, each image was graded for \$1.10, and grading garnered an effective hourly wage of \$14.31 (Table 2).

Overall number of correctly classified images decreased with the addition of a third and fourth grading category to 64.4% (148/230) and 50.9% (117/230), respectively. Specificity and sensitivity for individual Turkers was 96.3% and 66.7% respectively for both three and four categories. At the level of Turker consensus, sensitivity was 100% for both three and four categories, and specificity was 71.4% and 100% for three and four categories, respectively. With additional grading criteria, the speed of grading decreased, as did the effective hourly wage. Average time to complete the three-category HITs was 51 seconds, for an effective hourly wage of \$7.08. Average time to complete the four-category HITs was 55 seconds, for an effective hourly wage of \$6.60 (Table 2).

Table 1. Turker grading of individual images^a.

Image #	Two-category rating			Three-category rating			Four-category rating		
	Expert rating	Correct diagnosis ^b , %	Turker consensus ^c	Expert rating	Correct diagnosis ^b , %	Turker consensus ^c	Expert rating	Correct diagnosis ^b , %	Turker consensus ^c
1	Nor	65	—	Nor	90	—	Nor	55	—
2	Ab	85	—	M/M	50	Sev	Mild	0	Sev
3	Nor	70	—	Nor	70	—	Nor	70	—
4	Nor	50	Ab	Nor	40	M/M	Nor	60	—
5	Nor	80	—	Nor	70	—	Nor	50	—
6	Ab	100	—	M/M	90	—	Mild	20	Mod
7	Ab	90	—	Severe	60	—	Sev	10	Mod
8	Nor	50	Ab	Sev	40	M/M	Nor	65	—
9	Ab	100	—	Sev	95	—	Sev	100	—
10	Ab	100	—	Sev	40	M/M	Sev	70	—
11	Ab	90	—	Sev	0	M/M	Sev	20	Mild
12	Nor	90	—	Nor	80	—	Nor	90	—
13	Ab	100	—	M/M	30	Sev	Mod	20	Sev
14	Ab	80	—	Sev	40	M/M	Sev	10	Mod
15	Nor	90	—	Nor	100	—	Nor	90	—
16	Ab	90	—	Sev	70	—	Sev	50	—
17	Ab	100	—	M/M	60	—	Mild	10	Mod
18	Ab	100	—	M/M	100	—	Mod	95	—
19	Ab	90	—	M/M	80	—	Mild	20	Mod
Correct, %		81.3	89.5		64.4	63.2		50.9	57.9
Sensitivity ^d , %		93.6	100.0		96.3	100.0		96.3	100.0
Specificity ^d , %		67.8	71.4		66.7	71.4		66.7	100.0

^aNor=Normal; Ab=Abnormal; M/M=Mild or Moderate; Sev=Severe; Mod=Moderate.

^bAt the level of the individual graders.

^cConsensus rating presented only if it differed from the expert rating.

^dCalculated for normal versus any disease level.

Table 2. Time to complete ratings (in seconds).

	Two-category rating	Three-category rating	Four-category rating	Four-category rating (improved training)	Four-category rating (increased approval)	Four-category rating (Master Graders) ^a
Mean time per HITs	25.16	50.87	54.52	50.98	38.79	44.14
95% CI	21.93-28.38	43.18-58.55	46.15-62.88	39.66-62.30	31.65-45.93	36.00-52.27
Hourly wage, \$	14.31	7.08	6.60	7.06	9.28	12.23
Cost per image, \$	1.10	1.10	1.10	1.10	1.10	1.95

^aMaster graders received US \$0.15 per image, plus a 30% Amazon commission for a total cost of US \$0.195/image.

Phase II

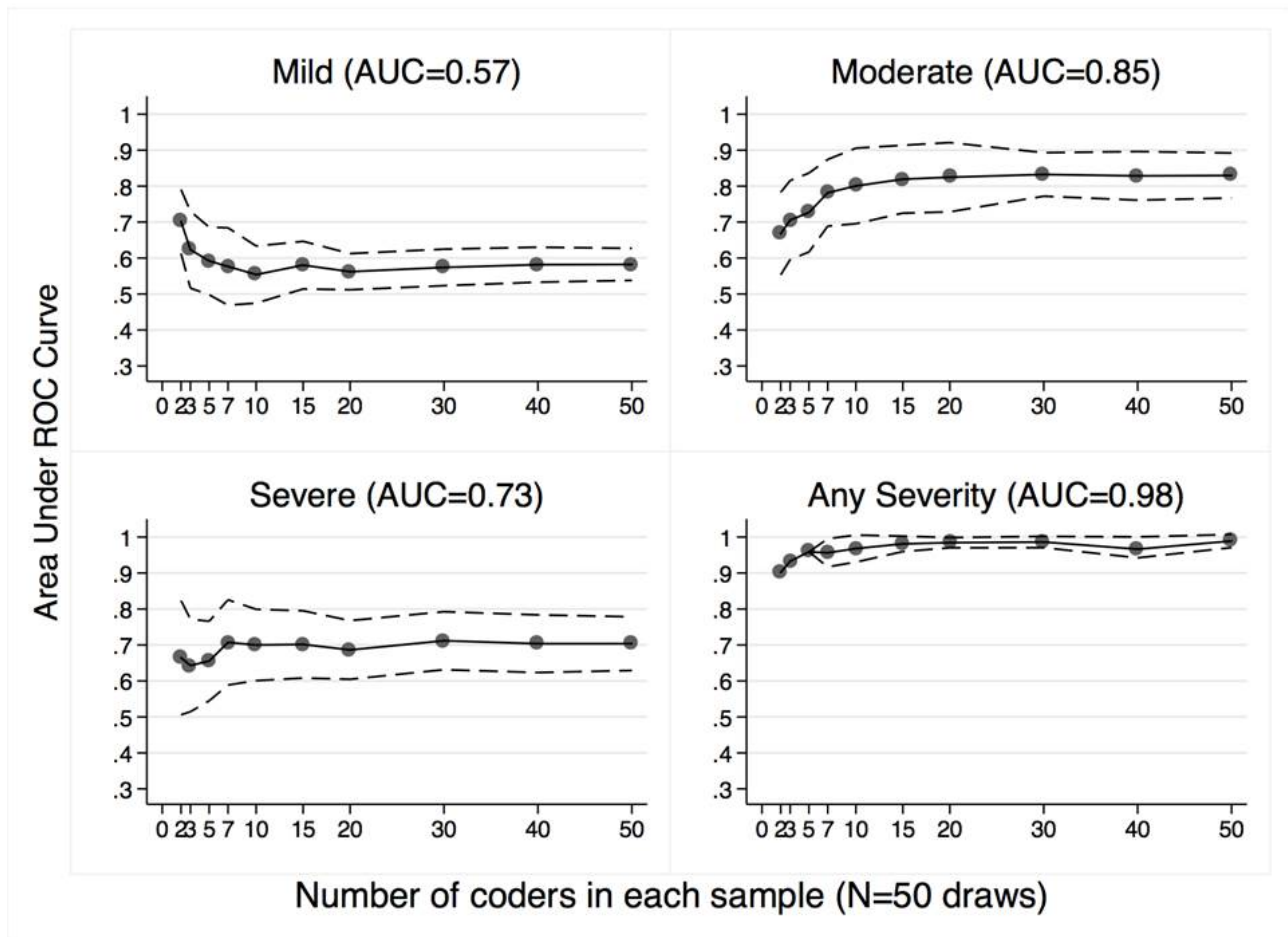
Results of Phase II likewise indicate that sensitivity and specificity for overall ratings of abnormal versus normal was excellent, producing a highly significant AUC (0.98; [Figure 2](#), Panel D). Turkers were not as accurate when asked to differentiate among four severity levels. Post hoc contrast

analyses, however, indicate that Turkers performed well when asked to identify the abnormalities that were moderate in severity (ROC=0.85; [Figure 2](#), Panel B). The pattern of results indicates that lower accuracy identifying mild (ROC=0.57; [Figure 2](#), Panel A) and severe (AUC=0.73; [Figure 2](#), Panel C) abnormalities was due to a tendency to rate all abnormalities as moderate in severity, rather than a failure to recognize normal

versus mild and severe abnormalities more generally. Results also indicate that maximum AUC was usually achieved when crowd size reached a total of between 7 and 10 Turkers, confirming the validity of the crowd sizes used to rate the larger

set of images (Figure 2). This affirms that the results of Phases I and III would not have been different had we sought a larger number of Turkers for each HIT.

Figure 2. Area under the curve (AUC) of the receiver-operator characteristic (ROC) curve for increasing numbers of Turker interpretations of a prototypical image from each severity level. Turkers had low accuracy for the Mild (Panel A) and Severe image (Panel C), but acceptable accuracy for the Moderate image (Panel B). When all four images were analyzed for absence or presence of disease only, Turkers performed well (Panel D) with a highly significant AUC.



Phase III

In Phase III, an additional normal training image was added due to Turkers’ interpreting visible choroidal vessels as abnormal during Phase I. The annotations were clarified to indicate that the presence of a single microaneurysm is considered abnormal and that hard exudates need not be present to achieve moderate or severe disease. Functionality to enlarge the image being graded was added. In the first batch using the new training, the percentage of correctly classified images using the consensus Turker scores was no better than previous: 42.1%

(8/19 images) versus 57.9% (11/19 images), respectively by mode and 21.1% (4/19 images) versus 26.3% (5/19 images) by mean. In the second batch, with more stringent criteria for Turker selection, the proportion of correctly graded images improved to 52.6% (10/19 images) by mean (mode unchanged). This result was confirmed in a repeat run of this exact batch (data not shown). In the third batch, using “Photo Moderation Masters”, the proportion of correctly graded images decreased to 36.8% (7/19 images) by mean (mode unchanged). Throughout all batches, the diagnostic sensitivity for any level of disease was maintained at 100% (Table 3)

Table 3. Turker consensus in Phase III.

	Number correct (mean) ^a	% correct (mean)	Number correct (mode) ^a	% correct (mode)	Sensitivity ^b	Specificity ^b
Phase I: Four-category rating	5	26.3	11	57.9	100.0	100.0
Phase 3: Trial 1 (improved training)	4	21.1 ^c	8 ^d	42.1	100.0	57.1
Phase 3: Trial 2 (raised approval rating)	10	52.6	11 ^e	57.9	100.0	100.0
Phase 3: Trial 3 (Master Graders)	7	36.8	11	57.9	100.0	100.0

^aCalculated by level (eg, Turker consensus matches expert designation as normal, mild, moderate, and severe).

^bCalculated for normal versus any disease level using the mode consensus score.

^cAfter excluding a single Turker with systematically higher scores, 42.1% correct.

^dThree images had no mode and were considered incorrect for “Number Correct” and “% correct” but recoded as abnormal for sensitivity and specificity.

^eOne image had no mode and was considered incorrect for “Number Correct” and “% correct” but recoded as abnormal for sensitivity and specificity.

Discussion

Principal Findings

With minimal training, an anonymous, untrained workforce recruited through a public crowdsourcing platform can rapidly and correctly categorize fundus photos of diabetic patients as normal or abnormal. The Turkers in this study graded these images with a high sensitivity, which is ideal for a screening instrument. Critically, no false negative consensus results were generated in any of our batches, indicating that no cases of DR were missed. When asked to categorize photos by degree of retinopathy, Turkers improved with changes in the Turk interface, particularly with increasing prior approval rating needed to perform the HITs. The number of graders required to reach a reliable “consensus grade” was consistent across categories, and 10 grading instances per image was established as sufficient for valid results.

Images were interpreted for a total cost of US \$1.10 per eye. While abnormal images would still currently still require further evaluation to identify patients in need of a live examination, this cost is near the limit suggested by some investigators for feasible public health screening in developing economies [34]. Indeed, the reimbursement per image, which ranged from an effective hourly wage of US \$6.60 to \$14.31 in our study, could possibly be reduced, since a reasonable target hourly wage for AMT workers is the federal minimum wage of US \$7.25 per hour or even lower [35]. Additionally, posting larger batches of images might allow for lower per image reimbursement, since Turker speed would likely increase after becoming more skilled at the task, allowing them to maintain the same effective wage. While there may not be a direct relationship between quality responses and high wages [36], there may be a link between high wages and rapid completion of image grading, so it may not be wise to dramatically reduce reimbursement.

A post hoc analysis of individual Turkers’ scores revealed inconsistent use of the four grading categories by one Turker in the first batch of Phase 3 (Table 3). Several issues are brought to light by considering this specific batch. First, inconsistent use of all categories was a rare occurrence, demonstrating that Turkers are conscientious workers. This was also evident from comments made by Turkers as they completed HITs, which

included thoughtful suggestions for improvement to the interface and concern over HITs that were felt to be ambiguous. Second, using the mean of crowdsourced responses may generate outputs that are rather sensitive to outliers. For this reason, using the mode to calculate consensus is generally preferable, though some images may not have a pure mode (Table 3), in which case the higher score of any “tie” would be used clinically. Third, and more broadly, AMT may be susceptible to Turker accounts that attempt to take advantage of the system by rapidly completing HITs with random responses either with live individuals or with automated programs or “bots” [29]. Moving forward, it may be necessary to analyze raw Turker scores for such phenomena and perhaps develop methods to exclude systematically unreliable scores.

Since AMT is a reputation-based economy, Requestors can reject or block Turkers who are not performing appropriately. Both actions negatively impact the Turkers’ reputation, which in turn affects their ability to perform HITs in the future, so there is a strong incentive to perform tasks accurately and honestly. This is likely why increasing the prior HIT approval rating to 99% had the most dramatic impact on consensus accuracy. Adding the “Photo Moderation Master” qualification did not improve worker consensus. This may be due to the fact that the criteria Amazon uses to grant this qualification are not relevant to our task. Additionally, since only a fraction of Turkers have the qualification, requiring it reduces the available workforce, which can increase the time required to complete batches. Especially when factoring in the additional Amazon commission, use of the Master qualification may not be necessary or cost-effective for these types of tasks in the future.

The current study was limited to a small set of hand-picked, high-quality, mydriatic fundus photos chosen to illustrate the key findings in diabetic retinopathy. Screening populations might have a subset of low-quality or uninterpretable images and would also be expected to have far more images of normal fundi. Identifying pathology in such sets would require extra vigilance on the part of Turkers to detect mild disease within large groups of normal photos. Larger datasets with more subtle pathology need to be tested with this methodology. Additionally, analyzing whether iterative improvements to the interface lead to better results is confounded by the fact that Turkers may have

previously been exposed to the task and may be improving in their grading through practice. This is unlikely because the Turkers receive no feedback on their grading, so they do not know if they have correctly graded images or not. Moreover, while it is not currently feasible to “block” Turkers who have previously completed any of our HITs, it is possible to view their grading history within the project. Surprisingly, throughout all batches posted, most of our HITs were completed by Turkers otherwise naïve to our project. In the final batch of 190 HITs posted for this project, after approximately 3000 HITs using the same images had been posted, 170 (89.5%) were completed by Turkers who had never done any of our HITs before, and 20 HITs were done by 3 individual Turkers who had each graded only four images previously (data not shown). In future larger batches, adjusting for individual graders’ improvement over time could become necessary.

Future Considerations

While further refinement of the methodology is still needed to resolve the degree of retinopathy, the current model could possibly be used as a way to reduce the burden on skilled graders by eliminating most of the normal images in very large screenings and passing only those abnormal images on for further characterization. While the individuals who complete HITs on AMT are truly anonymous, they do have unique numerical identifiers and can be tracked across HITs and batches of HITs. Therefore, an intriguing possibility using a crowdsourcing interface could include developing a cadre of specific Turkers who demonstrate a track record of reliable grading. These graders might be capable of a higher level of categorization than the general pool of Turkers and could be

recruited for more complex grading tasks. Additionally, it is likely that automated computer algorithms will also play a role in the analysis of fundus images for DR and other conditions in the future. This raises the possibility of an even more robust interaction between artificial intelligence and human intelligence. Images could be graded in a tiered fashion by one system, and then those graded ambiguously could be routed to the other for further validation.

An unanticipated benefit of such a crowdsourcing program is that it might raise awareness of diabetes and diabetic retinopathy. Since our interface allowed Turkers to leave feedback for us to refine the instrument, we were able to capture comments such as, “I have learn about diabetes little bit [sic]”, “I really liked seeing the pics of the eye, very interesting”, and “This HIT was very good and a nice break from all of the bubbling surveys. Thank you!”, suggesting an interest in the subject matter beyond other HITs and beyond what we had expected at the outset. This finding is consistent with what has been termed “Virtual Citizen Science” in fields outside of biomedical research [37].

Conclusions

Crowdsourcing represents a novel and inexpensive means to rapidly identify diabetic retinopathy. Further refinements of the technique are required, as is external validation with larger image sets. Additionally, multiple medico-legal and ethical issues would need to be addressed prior to clinical use of this technology, but there may be a role for crowdsourcing medical imaging data in large public health screenings and other settings in the future.

Acknowledgments

This research was supported by the J. Arch McNamara, MD Memorial Fund for Retina Research and Education at Wills Eye Hospital. The sponsor had no role in the design, conduct, or decision to publish the research, nor did they have any role in the review or approval of the manuscript.

This manuscript was presented at the American Society of Retinal Specialists Annual Meeting, August 2014.

Conflicts of Interest

None declared.

References

1. Beck RW. The burgeoning public health impact of diabetes: the role of the ophthalmologist. *Arch Ophthalmol* 2011 Feb;129(2):225-229. [doi: [10.1001/archophthalmol.2010.331](https://doi.org/10.1001/archophthalmol.2010.331)] [Medline: [21320972](https://pubmed.ncbi.nlm.nih.gov/21320972/)]
2. Javitt JC, Canner JK, Frank RG, Steinwachs DM, Sommer A. Detecting and treating retinopathy in patients with type I diabetes mellitus. A health policy model. *Ophthalmology* 1990 Apr;97(4):483-94; discussion 494. [Medline: [2109299](https://pubmed.ncbi.nlm.nih.gov/2109299/)]
3. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Ann Intern Med* 1996 Jan 1;124(1 Pt 2):164-169. [Medline: [8554212](https://pubmed.ncbi.nlm.nih.gov/8554212/)]
4. Jones S, Edwards RT. Diabetic retinopathy screening: a systematic review of the economic evidence. *Diabet Med* 2010 Mar;27(3):249-256. [doi: [10.1111/j.1464-5491.2009.02870.x](https://doi.org/10.1111/j.1464-5491.2009.02870.x)] [Medline: [20536486](https://pubmed.ncbi.nlm.nih.gov/20536486/)]
5. Schoenfeld ER, Greene JM, Wu SY, Leske MC. Patterns of adherence to diabetes vision care guidelines: baseline findings from the Diabetic Retinopathy Awareness Program. *Ophthalmology* 2001 Mar;108(3):563-571. [Medline: [11237912](https://pubmed.ncbi.nlm.nih.gov/11237912/)]
6. Paz SH, Varma R, Klein R, Wu J, Azen SP, Los Angeles Latino Eye Study Group. Noncompliance with vision care guidelines in Latinos with type 2 diabetes mellitus: the Los Angeles Latino Eye Study. *Ophthalmology* 2006 Aug;113(8):1372-1377. [doi: [10.1016/j.ophtha.2006.04.018](https://doi.org/10.1016/j.ophtha.2006.04.018)] [Medline: [16769120](https://pubmed.ncbi.nlm.nih.gov/16769120/)]

7. Lee PP, Feldman ZW, Ostermann J, Brown DS, Sloan FA. Longitudinal rates of annual eye examinations of persons with diabetes and chronic eye diseases. *Ophthalmology* 2003 Oct;110(10):1952-1959. [doi: [10.1016/S0161-6420\(03\)00817-0](https://doi.org/10.1016/S0161-6420(03)00817-0)] [Medline: [14522771](https://pubmed.ncbi.nlm.nih.gov/14522771/)]
8. Fong DS, Sharza M, Chen W, Paschal JF, Ariyasu RG, Lee PP. Vision loss among diabetics in a group model Health Maintenance Organization (HMO). *Am J Ophthalmol* 2002 Feb;133(2):236-241. [Medline: [11812428](https://pubmed.ncbi.nlm.nih.gov/11812428/)]
9. Sharp PF, Olson J, Strachan F, Hipwell J, Ludbrook A, O'Donnell M, et al. The value of digital imaging in diabetic retinopathy. *Health Technol Assess* 2003;7(30):1-119 [FREE Full text] [Medline: [14604499](https://pubmed.ncbi.nlm.nih.gov/14604499/)]
10. Scanlon PH. The English national screening programme for sight-threatening diabetic retinopathy. *J Med Screen* 2008;15(1):1-4. [doi: [10.1258/jms.2008.008015](https://doi.org/10.1258/jms.2008.008015)] [Medline: [18416946](https://pubmed.ncbi.nlm.nih.gov/18416946/)]
11. Ng M, Nathoo N, Rudnisky CJ, Tennant MT. Improving access to eye care: teleophthalmology in Alberta, Canada. *J Diabetes Sci Technol* 2009 Mar;3(2):289-296 [FREE Full text] [Medline: [20144360](https://pubmed.ncbi.nlm.nih.gov/20144360/)]
12. Teng T, Lefley M, Claremont D. Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy. *Med Biol Eng Comput* 2002 Jan;40(1):2-13. [Medline: [11954703](https://pubmed.ncbi.nlm.nih.gov/11954703/)]
13. Abramoff M, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol* 2013 Mar;131(3):351-357. [doi: [10.1001/jamaophthalmol.2013.1743](https://doi.org/10.1001/jamaophthalmol.2013.1743)] [Medline: [23494039](https://pubmed.ncbi.nlm.nih.gov/23494039/)]
14. Trucco E, Ruggeri A, Karnowski T, Giancardo L, Chaum E, Hubschman JP, et al. Validating retinal fundus image analysis algorithms: issues and a proposal. *Invest Ophthalmol Vis Sci* 2013 May;54(5):3546-3559 [FREE Full text] [doi: [10.1167/iovs.12-10347](https://doi.org/10.1167/iovs.12-10347)] [Medline: [23794433](https://pubmed.ncbi.nlm.nih.gov/23794433/)]
15. Brabham DC. *Crowdsourcing. The MIT Press Essential Knowledge Series.* Cambridge, MA: MIT Press; 2013.
16. Brabham DC, Ribisl KM, Kirchner TR, Bernhardt JM. Crowdsourcing applications for public health. *Am J Prev Med* 2014 Feb;46(2):179-187. [doi: [10.1016/j.amepre.2013.10.016](https://doi.org/10.1016/j.amepre.2013.10.016)] [Medline: [24439353](https://pubmed.ncbi.nlm.nih.gov/24439353/)]
17. Luengo-Oroz MA, Arranz A, Frea J. Crowdsourcing malaria parasite quantification: an online game for analyzing images of infected thick blood smears. *J Med Internet Res* 2012;14(6):e167 [FREE Full text] [doi: [10.2196/jmir.2338](https://doi.org/10.2196/jmir.2338)] [Medline: [23196001](https://pubmed.ncbi.nlm.nih.gov/23196001/)]
18. Mityr D, Peto T, Hayat S, Morgan JE, Khaw KT, Foster PJ. Crowdsourcing as a novel technique for retinal fundus photography classification: analysis of images in the EPIC Norfolk cohort on behalf of the UK Biobank Eye and Vision Consortium. *PLoS One* 2013;8(8):e71154 [FREE Full text] [doi: [10.1371/journal.pone.0071154](https://doi.org/10.1371/journal.pone.0071154)] [Medline: [23990935](https://pubmed.ncbi.nlm.nih.gov/23990935/)]
19. Li HK, Horton M, Bursell SE, Cavallerano J, Zimmer-Galler I, Tennant M, American Telemedicine Association Diabetic Retinopathy Telehealth Practice Recommendations Working Group, et al. Telehealth practice recommendations for diabetic retinopathy, second edition. *Telemed J E Health* 2011 Dec;17(10):814-837. [doi: [10.1089/tmj.2011.0075](https://doi.org/10.1089/tmj.2011.0075)] [Medline: [21970573](https://pubmed.ncbi.nlm.nih.gov/21970573/)]
20. Amazon Mechanical Turk. URL: <https://www.mturk.com/mturk/welcome> [accessed 2014-08-25] [WebCite Cache ID [6S51YMqkr](https://www.webcitation.org/6S51YMqkr)]
21. Ilakkuvan V, Tcelosky M, Ivey KC, Pearson JL, Cantrell J, Vallone DM, et al. Cameras for Public Health Surveillance: A Methods Protocol for Crowdsourced Annotation of Point-of-Sale Photographs. *JMIR Res Protoc* 2014;3(2):e22 [FREE Full text] [doi: [10.2196/resprot.3277](https://doi.org/10.2196/resprot.3277)] [Medline: [24717168](https://pubmed.ncbi.nlm.nih.gov/24717168/)]
22. Turner AM, Kirchoff K, Capurro D. Using crowdsourcing technology for testing multilingual public health promotion materials. *J Med Internet Res* 2012;14(3):e79 [FREE Full text] [doi: [10.2196/jmir.2063](https://doi.org/10.2196/jmir.2063)] [Medline: [22664384](https://pubmed.ncbi.nlm.nih.gov/22664384/)]
23. Henshaw EJ. Too sick, not sick enough? Effects of treatment type and timing on depression stigma. *J Nerv Ment Dis* 2014 Apr;202(4):292-299. [doi: [10.1097/NMD.0000000000000121](https://doi.org/10.1097/NMD.0000000000000121)] [Medline: [24647215](https://pubmed.ncbi.nlm.nih.gov/24647215/)]
24. Halkjelsvik T. Do disgusting and fearful anti-smoking advertisements increase or decrease support for tobacco control policies? *Int J Drug Policy* 2014 Jul;25(4):744-747. [doi: [10.1016/j.drugpo.2014.01.021](https://doi.org/10.1016/j.drugpo.2014.01.021)] [Medline: [24630079](https://pubmed.ncbi.nlm.nih.gov/24630079/)]
25. Carter RR, DiFeo A, Bogie K, Zhang GQ, Sun J. Crowdsourcing awareness: exploration of the ovarian cancer knowledge gap through Amazon Mechanical Turk. *PLoS One* 2014;9(1):e85508 [FREE Full text] [doi: [10.1371/journal.pone.0085508](https://doi.org/10.1371/journal.pone.0085508)] [Medline: [24465580](https://pubmed.ncbi.nlm.nih.gov/24465580/)]
26. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J Med Internet Res* 2012;14(2):e46 [FREE Full text] [doi: [10.2196/jmir.1988](https://doi.org/10.2196/jmir.1988)] [Medline: [22397809](https://pubmed.ncbi.nlm.nih.gov/22397809/)]
27. Good BM, Loguercio S, Griffith OL, Nanis M, Wu C, Su AI. The Cure: Design and Evaluation of a Crowdsourcing Game for Gene Selection for Breast Cancer Survival Prediction. *JMIR Serious Games* 2014 Jul 29;2(2):e7. [doi: [10.2196/games.3350](https://doi.org/10.2196/games.3350)]
28. Dasgupta N, Freifeld C, Brownstein JS, Menone CM, Surratt HL, Poppish L, et al. Crowdsourcing black market prices for prescription opioids. *J Med Internet Res* 2013;15(8):e178 [FREE Full text] [doi: [10.2196/jmir.2810](https://doi.org/10.2196/jmir.2810)] [Medline: [23956042](https://pubmed.ncbi.nlm.nih.gov/23956042/)]
29. Martin D, Hanrahan BV, O'Neill J. Being a turker. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. 2014 Presented at: 17th ACM conference on Computer supported cooperative work & social computing; February 15-19, 2014; Baltimore, Maryland, USA. [doi: [10.1145/2531602.2531663](https://doi.org/10.1145/2531602.2531663)]
30. Ipeirotis PG. CeDER Working Papers-10-01. 2010. Demographics of mechanical turk URL: <http://hdl.handle.net/2451/29585> [accessed 2014-09-26] [WebCite Cache ID [6StCw13A8](https://www.webcitation.org/6StCw13A8)]

31. Shotliff K, Duncan G. Diabetic retinopathy: summary of grading and management criteria. *Pract Diab Int* 2006 Nov;23(9):418-420. [doi: [10.1002/pdi.1029](https://doi.org/10.1002/pdi.1029)]
32. Quenouille MH. Notes on bias in estimation. *Biometrika* 1956;43(3-4):353-360. [doi: [10.1093/biomet/43.3-4.353](https://doi.org/10.1093/biomet/43.3-4.353)]
33. Amazon Mechanical Turk Frequently Asked Questions. URL: <https://requester.mturk.com/help/faq> [accessed 2014-08-25] [[WebCite Cache ID 6S5ljfanT](#)]
34. Friedman DS, Ali F, Kourgialis N. Diabetic retinopathy in the developing world: how to approach identifying and treating underserved populations. *Am J Ophthalmol* 2011 Feb;151(2):192-4.e1. [doi: [10.1016/j.ajo.2010.10.014](https://doi.org/10.1016/j.ajo.2010.10.014)] [Medline: [21251492](https://pubmed.ncbi.nlm.nih.gov/21251492/)]
35. Horton JJ, Chilton LB. The labor economics of paid crowdsourcing. In: Proceedings of the 11th ACM conference on Electronic commerce. 2010 Presented at: 11th ACM conference on Electronic commerce; June 7-11, 2010; Cambridge, Massachusetts, USA. [doi: [10.1145/1807342.1807376](https://doi.org/10.1145/1807342.1807376)]
36. Mason W, Watts DJ. Financial incentives and the "performance of crowds". In: Proceedings of the ACM SIGKDD Workshop on Human Computation. 2009 Presented at: 2009 ACM SIGKDD Workshop on Human Computation; June 28-July 1, 2009; Paris, France. [doi: [10.1145/1600150.1600175](https://doi.org/10.1145/1600150.1600175)]
37. Reed J, Raddick MJ, Lardner A. An exploratory factor analysis of motivations for participating in Zooniverse, a collection of virtual citizen science projects. In: 46th Hawaii International Conference on System Sciences (HICSS). 2013 Presented at: 46th Hawaii International Conference on System Sciences (HICSS); January 7-10, 2013; Wailea, Maui, HI. [doi: [10.1109/HICSS.2013.85](https://doi.org/10.1109/HICSS.2013.85)]

Abbreviations

AMT: Amazon Mechanical Turk
AUC: area under the curve
DR: diabetic retinopathy
HIT: human intelligence task
ROC: receiver-operator characteristic

Edited by G Eysenbach; submitted 25.08.14; peer-reviewed by D Brabham, G Leontidis; comments to author 10.09.14; revised version received 15.09.14; accepted 16.09.14; published 30.10.14.

Please cite as:

Brady CJ, Villanti AC, Pearson JL, Kirchner TR, Gupta OP, Shah CP
Rapid Grading of Fundus Photographs for Diabetic Retinopathy Using Crowdsourcing
J Med Internet Res 2014;16(10):e233
URL: <http://www.jmir.org/2014/10/e233/>
doi: [10.2196/jmir.3807](https://doi.org/10.2196/jmir.3807)
PMID: [25356929](https://pubmed.ncbi.nlm.nih.gov/25356929/)

©Christopher J Brady, Andrea C Villanti, Jennifer L Pearson, Thomas R Kirchner, Omesh P Gupta, Chirag P Shah. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 30.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Identifying Problematic Internet Users: Development and Validation of the Internet Motive Questionnaire for Adolescents (IMQ-A)

Christina Bischof-Kastner¹, Dipl Psych; Emmanuel Kuntsche², PhD; Jörg Wolstein¹, MD

¹Institute of Psychology, Faculty of Humanities, University of Bamberg, Bamberg, Germany

²Addiction Switzerland, Lausanne, Switzerland

Corresponding Author:

Jörg Wolstein, MD

Institute of Psychology

Faculty of Humanities

University of Bamberg

Markusstraße 8a

Bamberg, 96047

Germany

Phone: 49 951 863 2045

Fax: 49 951 863 4801

Email: joerg.wolstein@uni-bamberg.de

Abstract

Background: Internationally, up to 15.1% of intensive Internet use among adolescents is dysfunctional. To provide a basis for early intervention and preventive measures, understanding the motives behind intensive Internet use is important.

Objective: This study aims to develop a questionnaire, the Internet Motive Questionnaire for Adolescents (IMQ-A), as a theory-based measurement for identifying the underlying motives for high-risk Internet use. More precisely, the aim was to confirm the 4-factor structure (ie, social, enhancement, coping, and conformity motives) as well as its construct and concurrent validity. Another aim was to identify the motivational differences between high-risk and low-risk Internet users.

Methods: A sample of 101 German adolescents (female: 52.5%, 53/101; age: mean 15.9, SD 1.3 years) was recruited. High-risk users (n=47) and low-risk users (n=54) were identified based on a screening measure for online addiction behavior in children and adolescents (Online-Suchtverhalten-Skala, OSV_{K-S}). Here, “high-risk” Internet use means use that exceeds the level of intensive Internet use (OSV_{K-S} sum score ≥ 7).

Results: The confirmatory factor analysis confirmed the IMQ-A’s 4-factor structure. A reliability analysis revealed good internal consistencies of the subscales (.71 up to .86). Moreover, regression analyses confirmed that the enhancement and coping motive groups significantly predicted high-risk Internet consumption and the OSV_{K-S} sum score. A mixed-model ANOVA confirmed that adolescents mainly access the Internet for social motives, followed by enhancement and coping motives, and that high-risk users access the Internet more frequently for coping and enhancement motives than low-risk users. Low-risk users were primarily motivated socially.

Conclusions: The IMQ-A enables the assessment of motives related to adolescent Internet use and thus the identification of populations at risk. The questionnaire enables the development of preventive measures or early intervention programs, especially dealing with internal motives of Internet consumption.

(*J Med Internet Res* 2014;16(10):e230) doi:[10.2196/jmir.3398](https://doi.org/10.2196/jmir.3398)

KEYWORDS

Internet; adolescents; questionnaires; validation; addictive behavior; statistical factor analysis

Introduction

Internet use among so called “digital natives” [1,2] is linked to all areas of life [3-7]. For example, every German household containing young people aged between 12 and 19 years is

equipped with a computer or laptop [5]. In addition, personal computers are no longer the most common way of accessing the Internet in Europe. There has been unequivocal growth in access to the Internet via handheld or portable devices (eg, touchpads and smartphones), showing that the Internet is now

accessible to everyone [3,5]. Therefore, we can assume the existence of a generation that has grown up with the latest technologies from a very young age [1,2] and that Internet use is an extremely widespread phenomenon. This situation can be clearly explained by the fact that the Internet is a convenient source of information, social contacts, education, shopping, and recreational activities [6-10] that simplifies everyday life.

The Internet also has a negative side. The inclusion of Internet Gaming Disorder in Section III of the *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) (*DSM-5*) [11] illustrates the relevance of Internet addiction. Furthermore, different studies in several countries show that for 1.2% to 15.1% of cases, the intensive Internet use of adolescents can be classified as problematic [12-17]. Initial results from longitudinal studies even give rise to the suspicion that the disorder is highly stable [18].

To provide a basis for early intervention and preventive measures, we need to understand how the frequent Internet use of approximately 8% of adolescents becomes a problematic or dysfunctional behavior. Regarding the relationship between the Internet and all areas of life [5-7] and the suggested *DSM-5* criteria for Internet addiction [11], the duration someone spends online does not appear to be a valid criterion. Thus, getting to know the motives behind adolescent Internet consumption is important.

Concerning motives for media use in general, McQuail [19,20] assumes 4 basic motives: information, personal identity, integration and social identity, and entertainment, the latter covering relaxation, emotional relief, recreation, and more. Recent research regarding the motivations of Internet use in particular found the existence of instrumental motives, such as information seeking and social interaction, as well as a relationship between personality types and Internet use [6,21-28]. For example, Amiel and Sargent developed the Internet Motives Questionnaire [21], which measures the 4 instrumental motive groups interpersonal/communication utility, entertainment utility, information utility, and convenience.

There are, however, a number of gaps in current research. First, not much is known about affectivity, which plays an important role in understanding problematic Internet use. In fact, there are instruments for measuring Internet addiction and instrumental motives for using the Internet. In 1998, for example, Young [29] developed the first instrument for measuring Internet addiction, the Internet Addiction Test (IAT), which was translated into various languages and validated in different samples [30-33]. Also, the previously mentioned Internet Motives Questionnaire assesses motives that are more instrumental as they target the affective function the Internet serves for adolescents.

According to the motivation model of McClelland [34], affective change is assumed to be the driving force behind human behavior. Hence, striving for positive affect and escaping or relieving a dysphoric mood is seen as the basis of motivational (Internet) behavior and thus, according to the *DSM-5*, a criterion for Internet use disorder [11].

Second, none of the studies were based on a theoretical framework defining the dimensionality of motives for Internet use a priori. One exception is the Motives for Online Gaming Questionnaire (MOGQ) [35]. The MOGQ is based on a theoretical approach, although this model was developed by observing online players. Thus, the questionnaire is specific to Internet gaming and, therefore, not applicable to Internet use motives in general.

One promising approach that includes affectivity and allows the classification of the motivation behind human behavior was proposed by Cox and Klinger [36,37]. The Motivational Model of Alcohol Use assumes that people display a certain behavior to achieve expected or desired effects [38-42]. Motives can be classified by means of the valence and the source of the expected affective change. The valence is either positive (ie, to increase positive feelings) or negative (ie, to decrease negative feelings); the source can be internal (eg, in respect to one's own bodily sensations) or external (eg, in respect to significant others). Crossing these 2 dimensions results in the 4 motive categories enhancement, coping, conformity, and social as shown in Table 1 [40-44].

Table 1. Classification of drinking motives based on the Motivational Model of Alcohol Use by Cox and Klinger [36,37].

Source	Positive valence	Negative valence
Internal	Enhancement motives	Coping motives
External	Social motives	Conformity motives

Based on this 4-dimensional model, Cooper [38] developed the Drinking Motive Questionnaire (DMQ-R; German, French, and Italian version by Kuntsche et al [39]), which comprises 5 items per dimension. Although originally developed to understand alcohol use, the DMQ-R with its 4 motivational factors has already been successfully adapted to several behaviors beyond alcohol use (eg, gambling [41], sexual risk-taking behavior [42], and listening to music [43]). Given the success of previous adaptation attempts, the fact that McQuail's [19,20] basic motives for using media in general are reflected in the 4-dimensional model of Cooper [38] and the current state of research that assumes social interactions and conformity as

motives for Internet consumption (as described previously), it is likely that Cox and Klinger's motivational model can be applied to Internet use.

Consequently, based on the DMQ-R and the motivational model by Cox and Klinger [36,37], the aim of this study is to develop the Internet Motive Questionnaire for Adolescents (IMQ-A) (see Multimedia Appendices 1 and 2) as a theory-driven conceptualization and to validate it in a sample of German adolescents in a Web-based data collection conducted in the summer of 2011. More specifically, the following hypotheses are tested: (1) The IMQ-A 4-factor model has an adequate fit,

(2) every subscale of the IMQ-A (coping, enhancement, social, and conformity motives) has at least satisfactory internal consistency, (3) based on evidence from studies on alcohol use [38,39] and gambling [41], we expect enhancement and coping to be associated with high-score Internet use, but not social or conformity motives, and (4) problematic and nonproblematic Internet users differ with regards to their motives for going online.

Methods

Pretest

To examine the usefulness and quality of the IMQ-A, 70 students (female: 71%, 50/70; age: mean 15.90, SD 0.89 years) from 2 schools in Bamberg (Bavaria, Germany) completed a paper-and-pencil questionnaire in January 2011. The students had to answer the 20 items of the initial IMQ-A as well as questions concerning the comprehensibility of the items and the usefulness of the questionnaire. This allowed duration, comprehensibility of the content, and linguistic matching with the target group to be examined. Following the evaluation, particular items were adapted based on feedback from the students.

Study Design

Data were collected in a Web-based survey from June to September 2011 using the free software package

onlineFragebogen [45]. Online questionnaires are usually characterized by good acceptance among adolescents, high data quality, and external validity, particularly when participants complete the questionnaire voluntarily and without temporal constraint [46,47].

The participants were recruited in the German districts of Thuringia, Saxony, and Bavaria through posters, a press release, Facebook, various sports and youth clubs in Bamberg (Bavaria), and broadcasts by local radio stations. Interested individuals were included in a mailing list and provided with an automatically generated personal link required to log in to the study's website. Strict confidentiality and anonymity were guaranteed. Participants had to answer every question before proceeding to the next page.

After completing all 8 pages, adolescents wishing to take part in a random prize draw for a chance to win 1 of 3 Apple iPod Shuffles were invited to enter their email address. The email address was saved separately from the information collected in the questionnaire.

Participants

In total, 107 adolescents participated in the study. However, 6 participants were excluded because they were outside of the defined age range (14 to 19 years). The final sample consisted of 101 adolescents (age: mean 15.85, SD 1.33 years). Characteristics of the sample are shown in [Table 2](#).

Table 2. Sample information concerning demographics and general aspects of Internet use (N=107).

Characteristics with prevailing response options	n (%)
Sex	
Female	53 (52.2)
Male	48 (47.8)
Occupation	
Student	88 (87.1)
Trainee	6 (5.9)
Job-hunting/apprenticeship-hunting	4 (4.0)
Student at university	2 (2.0)
Employed part-time	1 (1.0)
Type of school/ educational qualifications	
High school	62 (61.4)
Secondary/junior high school	38 (37.6)
Other	1 (1.0)
Type of Internet access ^a	
Own computer/laptop	80 (79.2)
Web-enabled mobile phone	34 (33.7)
Parental computer/laptop	21 (20.8)
School computer/laptop	12 (11.9)
Parental control of Internet use	
Yes	18 (17.8)
No	83 (82.2)
Online frequency	
Daily	90 (89.1)
2-3 times per week	10 (9.9)
Once a week	1 (1.0)
Online activities ^a (very often or often)	
Entertainment (eg, music, videos, pictures)	96 (95.1)
Online communities (eg, Facebook)	94 (93.1)
Information research	86 (85.1)
Messenger (eg, ICQ, Skype)	59 (58.4)
Email contact	53 (52.5)
Chatting (eg, chat forums)	47 (46.6)
Shopping	21 (20.8)
Online games (eg, Wow, strategy games)	20 (19.8)
Online banking	4 (4.0)
Online gambling	0 (0.0)
Online sex offers	0 (0.0)

^a Multiple answers possible.

Measures

Demographics and General Aspects of Internet Use

This questionnaire contains information about age and the characteristics mentioned in [Table 2](#).

Screening Measure for Online Addiction Behavior in Children and Adolescents

The Online-Suchtverhalten-Skala (OSV_{K-S}) [48] (also Wölfling K, Müller KW, Beutel ME, unpublished data, 2009) used in this study is a German self-report measure for recording the addiction-like use of various Internet applications. This questionnaire was designed based on the screening scale for computer gaming behavior (Screeener zum Computerspielverhalten; CSV-R) [49] and the international classification criteria (*DSM-IV* [50]; ICD-10 [51]) for substance addictions [9] (also Wölfling K et al, unpublished data, 2009). The 16 items of the diagnostic module inquire about usage frequency, potential negative consequences of Internet use and pathological consumption patterns. With the help of 14 of these diagnostic items, a score can be computed that classifies the online behavior into the categories low-key, intensive, problematic/abusive, and addicted/pathological. In this case, however, only a subdivision into high-risk and low-risk Internet users was used. Here the risk for being prone to Internet addiction is meant. Therefore, in each case 2 categories were rolled into 1 (low-key and intensive to low-risk; problematic and addicted to high risk). Consequently, a cut-off point of approximately 7 points was used for group assignment.

Owing to the satisfactory to excellent results for psychometric characteristics, such as reliability ($\alpha=.88$), construct and factorial validity and clinical usefulness, the OSV_{K-S} is a promising instrument for Germany [52] (also Wölfling K et al, unpublished data, 2009).

Internet Motive Questionnaire for Adolescents

Based on the Drinking Motive Questionnaire Revised (DMQ-R) [38,39], the introduction is worded as follows: "Think of all the times you have been online during the last 12 months; how often do you go online..." Subsequently, 20 items measuring 4 dimensions were presented. Assuming the transferability of the statements, 12 items were taken from the DMQ-R without rewording (eg, "to forget your worries"); 8 items were adapted with regard to Internet use (eg, "because it improves parties and celebrations" was changed to "to improve contact with friends and acquaintances"). Each of the 5 items per dimension (ie, enhancement, coping, social, and conformity) were rated on a 5-point relative frequency scale with answer categories ranging from "(almost) never" (coded as 1) to "(almost) always" (coded as 5).

Statistical Analyses

Because some of the items of IMQ-A were adapted from a drinking motive questionnaire, we first conducted an explorative factor analysis using SPSS version 20 (IBM Corp, Armonk, NY, USA).

Confirmatory factor analysis (CFA) was used to confirm the 4-factor structure and construct validity of the 16-item IMQ-A. The CFA was performed with SPSS Amos version 20 using the Bollen-Stine bootstrap correction to account for nonnormal distribution [53]. Errors were allowed to correlate. To evaluate the model fit, the fit indices chi-square divided by degrees of freedom (reduced chi-square, χ^2_{red}), comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean squared residual (SRMR) were used. The parsimony measurement χ^2_{red} suggests an acceptable fit with values between 1 and 2 [54], whereas CFI and TLI should show values greater than .90 [55]. Regarding RMSEA and SRMR, results lower than .10 were sought; values ranging from .05 to .08 were treated as acceptable fit, and figures between .08 and .10 as moderate fit [55,56].

Cronbach alpha values were used as a measure of internal consistency for which values greater or equal to .9, .8, and .7 are considered as excellent, good, and acceptable, respectively [53,57].

To test concurrent validity of the IMQ-A, multiple regression was performed with the 4 motive groups as independent variables. Due to the dichotomy of the dependent variable high-risk Internet use, a logistic regression was performed. In the second model, a multiple linear regression, the OSV_{K-S} sum score was the dependent variable.

Differences in the 4 motive dimensions in the entire sample and among high-risk and low-risk users were tested using a 2 (high-risk vs low-risk users) \times 4 (motive dimension) mixed-model analysis of variance (ANOVA). Whenever an overall effect or an interaction was significant, post hoc tests (Bonferroni) were conducted to determine whether high-risk or low-risk users differed on a motive dimension. Descriptive analyses and the mixed-model ANOVA were performed using the SPSS 20.0 statistical software package.

Results

Descriptive Statistics

Descriptive results show that 47 of 107 participants (46.5%) were identified as high-risk users and 54 (53.5%) as low-risk users. The groups did not differ significantly in age, but did in gender and education ([Table 3](#)). Most high-risk users were male (32/47, 68%), and high school and secondary school students were equally represented (18/47, 38% each). By contrast, the low-risk users were mostly female (38/54, 70%) and grammar school students (44/54, 81%). Significant differences between the 2 groups were also found in terms of variables regarding the Internet behavior of the respondents (eg, online duration) ([Table 3](#)).

Although no significant differences emerged regarding the duration of use on weekends, online frequency, and parental control, a higher online duration on weekdays was reported by high-risk users. This pattern also became apparent when examining the regular online duration (high risk: mean 2.63, SD 1.61 hours; low risk: mean 1.77, SD 0.96 hours).

Table 3. Sociodemographic- and consumption-relevant characteristics of low-risk and high-risk adolescent Internet users (N=107).

Characteristics	Low risk (n=54)	High risk (n=47)	χ^2_1	t_{99}	U	P
Male, n (%)	16 (29.6)	32 (68.1)	14.9			<.001
Parental control, n (%)	8 (14.8)	10 (21.3)	0.7			.40
Age, mean (SD)	15.94 (1.31)	15.74 (1.36)		0.752		.45
Internet use on weekdays (in hours), mean (SD)	2.63 (2.40)	5.00 (4.14)		-2.685		.009
Internet use at the weekend/on holidays (in hours), mean (SD)	3.37 (1.83)	3.68 (2.60)		-0.482		.63
Education, mean rank	40.47	63.10			700.5	<.001
Online frequency, mean rank	52.08	49.76			1210.5	.46
Online duration, mean rank	44.54	58.43			920.0	.01
Online activity "information research", mean rank	55.78	45.51			1011.0	.046

Confirming the Four-Factor Structure

To ensure that each item loaded on the dimension to which it theoretically belonged, we first conducted an explorative factor analysis (principal axes factor analysis, varimax rotation with Kaiser normalization). The Kaiser-Meyer-Olkin measure of sampling adequacy was .815; the Bartlett's test of sphericity was significant. The results of the principal axes factor analysis showed the items "because it helps you enjoy your free time,"

"so that others don't make fun of you," "to be in high spirits," and "to have more self-confidence" were not strongly related to any dimension. In addition, leaving out these items from the dimension to which they were intended to belong resulted in higher explained variances of the factors and thus in total (Table 4). Consequently, only 16 items (4×4 solution) were included in the final version of the IMQ-A. The wording of all included items is provided in Table 5.

Table 4. Factor loadings of the 4 omitted items and coefficients of determination (R^2) of the factors from principal axes analysis following varimax rotation (n=70).

Item	Factor 1: coping	Factor 2: social	Factor 3: enhancement	Factor 4: conformity
Because it helps you to enjoy your free time	.329	-.344	.389	.229
So that others don't make fun of you		-.336		.360
To be in high spirits	.591	-.293	.412	
To have more self-confidence	.578	-.413		.370
R^2 including the items	29.2	13.6	9.6	6.7
R^2 excluding the items	29.0	15.6	10.2	8.1

The given identifiability (number of estimated parameters is lower than the number of sample moments; $40 < 136$) justifies the performance of a specified CFA [54]. The CFA of the 16-item IMQ-A yielded significant factor loadings ranging from $\lambda=.22$ to $\lambda=.89$ (Table 5). The factor coping had the highest item loadings, followed by the factor social. Consequently, these 2 factors also showed the highest internal consistencies. The factor conformity had the lowest loadings and internal

consistency. The highest correlation was found between enhancement and coping, the lowest between coping and social. The fit indexes χ^2_{red} , CFI, and RMSEA suggested good model fit ($\chi^2_{red}=1.5$; CFI=.912; RMSEA=.071, 95% CI .045-.094). TLI and SRMR were very close to the recommended thresholds (TLI=.890; SRMR=.088); therefore, they indicated a merely acceptable fit [54-56].

Table 5. Item factor loadings, item means, interfactor correlations, and internal consistencies as results of the confirmatory factor analysis to test the 4-factor structure of motives for Internet use.

Items of the IMQ-A	Enhancement	Coping	Social	Conformity	<i>P</i>	Mean (SD)
How often do you go online (<i>r</i>):						
Because it gives you a pleasant feeling?	.77				<.001	2.60 (1.21)
Because it is exciting?	.63				<.001	2.35 (1.13)
To experience a feeling of exaltation?	.69				<.001	1.68 (0.88)
Simply because it is fun?	.29				.01	3.76 (0.97)
To forget your worries?		.77			<.001	2.33 (1.14)
Because it helps you when you feel depressed or irritated?		.74			<.001	2.72 (1.13)
To cheer yourself up when you are in a bad mood?		.66			<.001	2.87 (1.11)
To forget about your problems?		.87			<.001	2.40 (1.20)
To come into contact with others?			.89		<.001	3.51 (1.35)
Because it is fun to be in contact with others?			.65		<.001	2.86 (1.18)
To improve your contact with friends and acquaintances?			.82		<.001	3.10 (1.15)
To share a special occasion with friends?			.55		<.001	2.76 (1.16)
Because your friends pressurized you to do it?				.22	.08	1.41 (0.70)
Because you would like to belong to a certain circle of friends?				.75	<.001	1.45 (0.79)
To be liked by others?				.38	.005	1.52 (0.81)
To not feel excluded?				.61	<.001	1.61 (0.94)
Interfactor correlations, <i>r</i> (<i>P</i>)						
Coping×	.66 (<.001)					
Social×	-.38 (.007)	-.52 (<.001)				
Conformity×	.32 (.05)	.30 (.04)	.10 (.47)			
Internal consistencies, Cronbach α	.80	.86	.82	.71		

Concurrent Validity of the IMQ-A

The logistic regression analysis revealed that enhancement and coping but not social and conformity were significantly related to high-risk Internet use (Table 6). The 4 motive dimensions explained approximately 50% of the variance in high-risk

Internet use ($R^2=49.6\%$). Testing the relationship between the 4 IMQ-A dimensions and the OSV_K-S score revealed an effect of enhancement and coping as well as conformity (Table 6), but not of social. The explained variance in the OSV_K-S score was 33.8%.

Table 6. Motives for Internet use as predictors of problematic Internet use and the OSV_K-S sum score.

Motives	Dysfunctional Internet use? (yes/no)		OSV _K -S sum score	
	OR (95% CI)	<i>P</i>	β^a	<i>P</i>
Social	0.85 (0.77, 1.01)	.06	-.16	.09
Enhancement	1.33 (1.08, 1.64)	.008	.29	.004
Coping	1.19 (1.02, 1.40)	.03	.22	.04
Conformity	1.17 (0.92, 1.48)	.21	.22	.02

^a Standardized regression weight (β).

Motive Ranking Orders

In the entire sample, adolescents most frequently accessed the Internet due to social motives followed by enhancement or

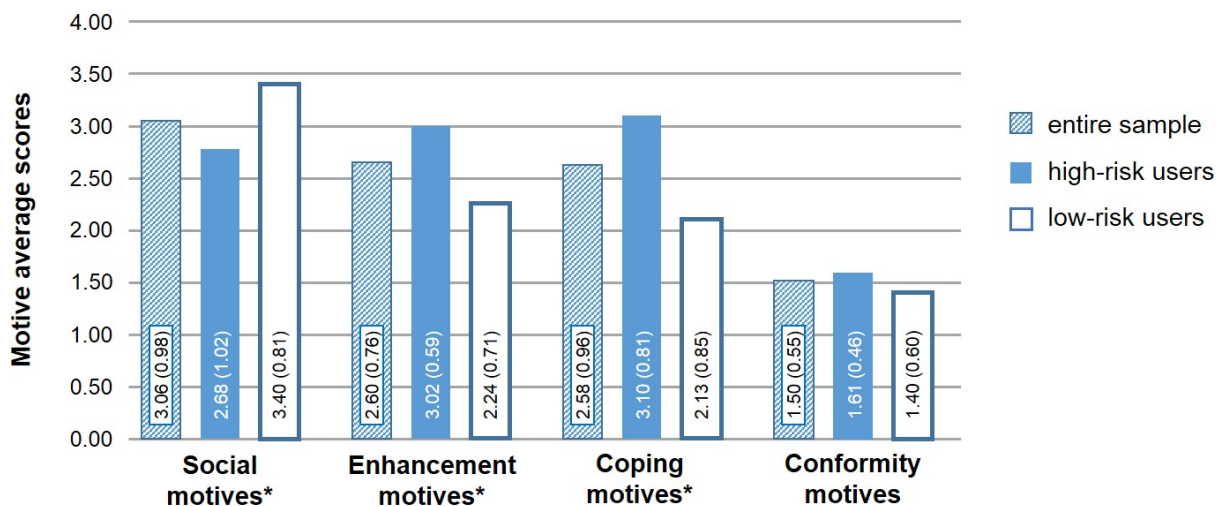
coping motives (Figure 1). Conformity motives rarely or never applied to participants.

Comparing the motive ranking order in both groups (high-risk and low-risk users), the 2×4 mixed-model ANOVA revealed a

group membership ($F_{1, 101}=15.317$; $P<.001$) main effect (high risk > low risk) as well as an Internet motive dimension ($F_{2,297, 101}=81.24$; $P<.001$) main effect (social > enhancement > coping > conformity). These effects were qualified by a motive \times group interaction ($F_{2,297, 101}=27.196$, $P<.001$). Post hoc (Bonferroni) tests revealed a significant difference for enhancement, coping,

and social, but not for conformity (Figure 1). Members of the high-risk group were predominantly motivated internally in their Internet consumption (coping > enhancement > social > conformity), whereas low-risk users indicated both internal and external motives, but were primarily motivated socially (social > enhancement > coping > conformity).

Figure 1. Comparison of the average scores of the 4 motive groups in the total sample and according to risk group. Shown are mean (SD). * indicates $P<0.001$.



Discussion

To provide a basis for early intervention and preventive measures, understanding the motives that drive intensive Internet use to become dysfunctional among adolescents is important. Therefore, the aim of this study was to develop the IMQ-A as a theory-based measurement for identifying the underlying motives for high-risk Internet use.

Because every adolescent uses the Internet daily [3,5-7] and using the Internet is now regarded as the fourth cultural competence, the widespread nature of the Internet is undisputed. Because of this, distinguishing between intensive and dysfunctional adolescent Internet use becomes difficult. Studies regarding 3 other problematic behaviors of adolescents show that motives are a predictor for addictive behavior. Furthermore, motives are comparable across different cultures [41-43,58], which could be explained by the link between motives and personality [22,24-27,58]. As a consequence, developing the IMQ-A based on a questionnaire that had already been developed and validated for different behaviors and cultures not only allows the underlying motives for high-risk Internet use to be identified, but it also offers starting points for developing preventive measures or early intervention programs as well as the chance to compare motivations for different behaviors across different cultures.

Testing the construct validity, the results of the CFA revealed an acceptable, but not outstanding, fit of the motives for Internet use from the 4-factor model. Two factors may be responsible for the comparatively low model fit. First, we used a rather small sample, which turned out to be quite heterogeneous because of the various different recruitment methods. Secondly,

IMQ-A items were developed based on an instrument constructed to measure drinking motives. Despite considerable similarities between Internet motives and drinking motives, the item “because your friends pressured you to do it,” for instance, seems to be worded too actively and too strongly for the Internet context, which is reflected by a weak or nonsignificant item loading. Nevertheless, all the other items showed significant loadings on the expected scales. Moreover, at least satisfactory internal consistencies were obtained for all dimensions [53].

The results of the regression analyses support the concurrent validity of the IMQ-A. In accordance with the literature concerning drinking motives [59,60] and gambling motives [41], enhancement and coping predicted the 2 criteria of dysfunctional Internet use (prediction of high-risk users and higher OSV_{K-S} score). The consistency of results across the 2 indicators highlights the robustness of the findings and indicates that accessing the Internet frequently for internal emotion-regulation motives is more dysfunctional than social Internet use, which appears more recreational. Our findings showed that persons characterized as low-risk users seem to use the Internet mainly to meet social needs and maintain contacts. Consequently, it can be supposed that the Internet (as is the case with alcohol consumption or gambling) is also used as a method for regulating emotions, and when this motivation comes into play, Internet consumption may exceed the useful aspects.

Concluding from the results of the CFA and the regression analyses, construct and concurrent validity can be regarded as given. Therefore, it can be assumed that motives of adolescent Internet consumption can be measured using the developed questionnaire. By implication, the respondents’ answers reveal that they use the Internet both for information research and

reasons of social identity and integrity as well as for entertainment as defined by McQuail [19,20].

The results regarding the motive ranking suggest that adolescents access the Internet primarily for positive motives (enhancement and social), whereas negative motives (coping and confirmatory) seem to have a minor impact. Interestingly, these results are consistent with the drinking motives literature [38,40,60,61]. Further analogies to the field of drinking motives can be drawn [40] based on the results of group-specific analysis. Thus, high-risk users mainly accessed the Internet for internal motives, whereas low-risk users indicated both internal and external motives, but were primarily motivated socially.

Although the presented results demonstrate the psychometric qualities of the IMQ-A to assess adolescent motives for accessing the Internet, it is likely that there are other motives for accessing the Internet besides affective change, such as gaming or knowledge acquisition [3-5,22-25]. Another limitation is the small sample size and the fact that the current sample is not representative of all Internet users in this age group. For example, compared with other studies [12-17], the prevalence of dysfunctional Internet users (46.5%) indicates an overrepresentation of problematic users. One explanation could

be the Web-based data collection method as well as the fact that 2 OSV_{K-S} categories were rolled into 1 (problematic and addicted to high risk).

Consequently, we recommend further validation of the IMQ-A with larger sample sizes, possibly including other cultures or parts of Germany. Moreover, in the presented survey, the IMQ-A was applied to a nonclinical sample. For this reason, it remains unclear to what extent the questionnaire can be used in a clinical sample. Future longitudinal studies are recommended to further examine the predictive validity. In this way, tests could be conducted to see whether the IMQ-A predicts future Internet use behavior among adolescents.

Despite the described limitations, the IMQ-A appears to be a valid and reliable instrument to assess motives related to adolescent Internet use. This questionnaire could serve as a basis on which to develop preventive measures or early intervention programs dealing especially with internal motives of Internet consumption.

In the clinical field, the IMQ-A can help to identify the motives of dysfunctional Internet consumption and establish individual intervention aspects for developing alternative coping strategies for coping users, for example.

Acknowledgments

The authors would like to thank all respondents for taking part in the study. Furthermore, the authors wish to express their gratitude to the Bavarian Academy for Addiction and Health (Bayerische Akademie für Sucht- und Gesundheitsfragen eV; BAS) for providing funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The Internet Motive Questionnaire for Adolescents (IMQ-A; German version).

[PDF File (Adobe PDF File), 59KB - [jmir_v16i10e230_app1.pdf](#)]

Multimedia Appendix 2

The Internet Motive Questionnaire for Adolescents (IMQ-A; English version).

[PDF File (Adobe PDF File), 59KB - [jmir_v16i10e230_app2.pdf](#)]

References

1. Prensky M. Digital natives, digital immigrants. In: Blair K, Almjeld J, Murphy RM, editors. *Cross Currents: Cultures, Communities, Technologies*. Boston: Wadsworth; 2006:45-51.
2. Joiner R, Gavin J, Brosnan M, Cromby J, Gregory H, Guiller J, et al. Comparing first and second generation digital natives' Internet use, Internet anxiety, and Internet identification. *Cyberpsychol Behav Soc Netw* 2013 Jul;16(7):549-552. [doi: [10.1089/cyber.2012.0526](#)] [Medline: [23675995](#)]
3. Livingstone S, Haddon L, Görzig A, Ólafsson K. Risks and safety on the internet: the perspective of European children. Full Findings. London: EU Kids Online; 2011. URL: <http://www.lse.ac.uk/media%40lse/research/EUKidsOnline/EU%20Kids%20II%20%282009-11%29/EUKidsOnlineIIReports/D4FullFindings.pdf> [accessed 2014-02-02] [WebCite Cache ID 6N5UJ51Tp]
4. comScore. 2012. The European digital day URL: <http://www.comscoredatamine.com/2012/09/the-european-digital-day/> [accessed 2014-02-02] [WebCite Cache ID 6N5XgFlhT]

5. Rathgeb T. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger in Deutschland [Baseline study for media exposure of 12- up to 19-year-old in Germany]. Deutschland: Jugend, Information, (Multi-)Media (JIM); 2012. URL: http://www.mpfs.de/fileadmin/JIM-pdf12/JIM2012_Endversion.pdf [accessed 2013-12-09] [WebCite Cache ID 6Ljz07uYs]
6. Bundesverband Informationswirtschaft, Telekommunikation und neue Medien eV. Jugend 2.0. Eine repräsentative Untersuchung zum Internetverhalten von 10- bis 18-Jährigen [A representative online behavior study concerning 10- up to 18-year-old]. Berlin: BITKOM; 2011. URL: http://www.bitkom.org/files/documents/BITKOM_Studie_Jugend_2.0.pdf [accessed 2013-12-09] [WebCite Cache ID 6LjzGjgcv]
7. Livingstone S. Children and the Internet. Cambridge: Polity; 2009.
8. Kuntsche E, Simons-Morton B, ter Bogt T, Queija IS, Tinoco VM, Gaspar de Matos M, HBSC Peer Culture Focus Group. Electronic media communication with friends from 2002 to 2006 and links to face-to-face contacts in adolescence: an HBSC study in 31 European and North American countries and regions. *Int J Public Health* 2009 Sep;54 Suppl 2:243-250 [FREE Full text] [doi: [10.1007/s00038-009-5416-6](https://doi.org/10.1007/s00038-009-5416-6)] [Medline: [19623474](https://pubmed.ncbi.nlm.nih.gov/19623474/)]
9. Müller KW, Wölfling K. Pathologische Computerspiel- und Internetnutzung - Der Forschungsstand zu Phänomenologie, Epidemiologie, Diagnostik und Komorbidität. *Suchtmedizin* 2010;12(01):45-55. [doi: [10.1055/s-0031-1275314](https://doi.org/10.1055/s-0031-1275314)]
10. Beutel ME, Brähler E, Glaesmer H, Kuss DJ, Wölfling K, Müller KW. Regular and problematic leisure-time Internet use in the community: results from a German population-based survey. *Cyberpsychol Behav Soc Netw* 2011 May;14(5):291-296. [doi: [10.1089/cyber.2010.0199](https://doi.org/10.1089/cyber.2010.0199)] [Medline: [21067277](https://pubmed.ncbi.nlm.nih.gov/21067277/)]
11. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Arlington, VA: Amer Psychiatric Pub; 2013.
12. Dreier M, Tzavela E, Wölfling K, Mavromati F, Duven E, Karakitsou C, et al. 1653 – The development of adaptive and maladaptive patterns of internet use among european adolescents at risk for internet addictive behaviors: a grounded theory inquiry. *Eur Psychiat* 2013 Jan;28(1):1-10. [doi: [10.1016/S0924-9338\(13\)76643-4](https://doi.org/10.1016/S0924-9338(13)76643-4)]
13. Cao F, Su L. Internet addiction among Chinese adolescents: prevalence and psychological features. *Child Care Health Dev* 2007 May;33(3):275-281. [doi: [10.1111/j.1365-2214.2006.00715.x](https://doi.org/10.1111/j.1365-2214.2006.00715.x)] [Medline: [17439441](https://pubmed.ncbi.nlm.nih.gov/17439441/)]
14. Vilella C, Martinotti G, Di Nicola M, Cassano M, La Torre G, Gliubizzi MD, et al. Behavioural addictions in adolescents and young adults: results from a prevalence study. *J Gambl Stud* 2011 Jun;27(2):203-214. [doi: [10.1007/s10899-010-9206-0](https://doi.org/10.1007/s10899-010-9206-0)] [Medline: [20559694](https://pubmed.ncbi.nlm.nih.gov/20559694/)]
15. Sasmaz T, Oner S, Kurt A, Yapici G, Yazici AE, Bugdayci R, et al. Prevalence and risk factors of Internet addiction in high school students. *Eur J Public Health* 2014 Feb;24(1):15-20. [doi: [10.1093/eurpub/ckt051](https://doi.org/10.1093/eurpub/ckt051)] [Medline: [23722862](https://pubmed.ncbi.nlm.nih.gov/23722862/)]
16. Hawi NS. Internet addiction among adolescents in Lebanon. *Comput Hum Behav* 2012 May;28(3):1044-1053. [doi: [10.1016/j.chb.2012.01.007](https://doi.org/10.1016/j.chb.2012.01.007)]
17. Siomos KE, Dafouli ED, Braimiotis DA, Mouzas OD, Angelopoulos NV. Internet addiction among Greek adolescent students. *Cyberpsychol Behav* 2008 Dec;11(6):653-657. [doi: [10.1089/cpb.2008.0088](https://doi.org/10.1089/cpb.2008.0088)] [Medline: [18991535](https://pubmed.ncbi.nlm.nih.gov/18991535/)]
18. Gentile DA, Choo H, Liau A, Sim T, Li D, Fung D, et al. Pathological video game use among youths: a two-year longitudinal study. *Pediatrics* 2011 Feb;127(2):319-329 [FREE Full text] [doi: [10.1542/peds.2010-1353](https://doi.org/10.1542/peds.2010-1353)] [Medline: [21242221](https://pubmed.ncbi.nlm.nih.gov/21242221/)]
19. Reimer M. Eskapismus-Konzept [Escapism concept]. In: Sander U, Gross F, Hugger KU, editors. *Handbuch Medienpädagogik*. Wiesbaden: Verlag für Sozialwissenschaften; 2008:297-300.
20. McQuail D. *McQuail's Mass Communication Theory*. London: Sage Publications Ltd; 2010.
21. Sargent SL. Internet Motives Questionnaire. In: Reynolds RA, Woods RA, Baker JD, editors. *Handbook of Research on Electronic Surveys and Measurements*. London: IGI Global; 2007:371-375.
22. Papacharissi Z, Rubin AM. Predictors of Internet use. *Journal Broadcast Electron* 2000 Jun;44(2):175-196. [doi: [10.1207/s15506878jobem4402_2](https://doi.org/10.1207/s15506878jobem4402_2)]
23. Flaherty LM, Pearce KJ, Rubin RB. Internet and face - to - face communication: not functional alternatives. *Comm Q* 1998 Jun;46(3):250-268. [doi: [10.1080/01463379809370100](https://doi.org/10.1080/01463379809370100)]
24. Swickert RJ, Hittner JB, Harris JL, Herring JA. Relationships among Internet use, personality, and social support. *Comp Hum Behav* 2002 Jul;18(4):437-451. [doi: [10.1016/S0747-5632\(01\)00054-1](https://doi.org/10.1016/S0747-5632(01)00054-1)]
25. Hamburger Y, Ben-Artzi E. The relationship between extraversion and neuroticism and the different uses of the Internet. *Comput Hum Behav* 2000 Jul;16(4):441-449. [doi: [10.1016/S0747-5632\(00\)00017-0](https://doi.org/10.1016/S0747-5632(00)00017-0)]
26. Tuten TL, Bosnjak M. Understanding differences in Web usage: the role of need for cognition and the five-factor model of personality. *Soc Behav Personal* 2001 Jan 01;29(4):391-398. [doi: [10.2224/sbp.2001.29.4.391](https://doi.org/10.2224/sbp.2001.29.4.391)]
27. Hetzel-Riggin MD, Pritchard JR. Predicting problematic Internet use in men and women: the contributions of psychological distress, coping style, and body esteem. *Cyberpsychol Behav Soc Netw* 2011 Sep;14(9):519-525. [doi: [10.1089/cyber.2010.0314](https://doi.org/10.1089/cyber.2010.0314)] [Medline: [21342011](https://pubmed.ncbi.nlm.nih.gov/21342011/)]
28. Choi J, Watt J, Dekkers A, Park SH. Motives of internet uses: Crosscultural perspective - the US, the Netherlands, and S Korea. 2004. URL: <http://hbo-kennisbank.uvt.nl/cgi/fontys/show.cgi?fid=49> [accessed 2013-12-09] [WebCite Cache ID 6Lk0DuVog]
29. Young KS. Internet addiction: the emergence of a new clinical disorder. *Cyberpsychol Behav* 1998 Jan;1(3):237-244. [doi: [10.1089/cpb.1998.1.237](https://doi.org/10.1089/cpb.1998.1.237)]

30. Widyanto L, McMurran M. The psychometric properties of the internet addiction test. *Cyberpsychol Behav* 2004 Aug;7(4):443-450. [doi: [10.1089/cpb.2004.7.443](https://doi.org/10.1089/cpb.2004.7.443)] [Medline: [15331031](https://pubmed.ncbi.nlm.nih.gov/15331031/)]
31. Khazaal Y, Billieux J, Thorens G, Khan R, Louati Y, Scarlatti E, et al. French validation of the internet addiction test. *Cyberpsychol Behav* 2008 Dec;11(6):703-706. [doi: [10.1089/cpb.2007.0249](https://doi.org/10.1089/cpb.2007.0249)] [Medline: [18954279](https://pubmed.ncbi.nlm.nih.gov/18954279/)]
32. Barke A, Nyenhuis N, Kröner-Herwig B. The German version of the internet addiction test: A validation study. *Cyberpsychol Behav Soc Netw* 2012;15(10):534-542. [doi: [10.1089/cyber.2011.0616](https://doi.org/10.1089/cyber.2011.0616)] [Medline: [23002984](https://pubmed.ncbi.nlm.nih.gov/23002984/)]
33. Faraci P, Craparo G, Messina R, Severino S. Internet Addiction Test (IAT): Which is the best factorial solution? *J Med Internet Res* 2013;15(10):e225 [FREE Full text] [doi: [10.2196/jmir.2935](https://doi.org/10.2196/jmir.2935)] [Medline: [24184961](https://pubmed.ncbi.nlm.nih.gov/24184961/)]
34. McClelland DC. *Human Motivation*. Cambridge, UK: Cambridge University Press; 1987.
35. Demetrovics Z, Urbán R, Nagygyörgy K, Farkas J, Zilahy D, Mervó B, et al. Why do you play? The development of the motives for online gaming questionnaire (MOGQ). *Behav Res Methods* 2011;43(3):814-825. [doi: [10.3758/s13428-011-0091-y](https://doi.org/10.3758/s13428-011-0091-y)] [Medline: [21487899](https://pubmed.ncbi.nlm.nih.gov/21487899/)]
36. Cox WM, Klinger E. A motivational model of alcohol use. *J Abnorm Psychol* 1988;97(2):168-180. [Medline: [3290306](https://pubmed.ncbi.nlm.nih.gov/3290306/)]
37. Cox WM, Klinger E. Incentive motivation, affective change in alcohol use: A model. In: *Why People Drink: Parameters of Alcohol as a Reinforcer*. New York: Gardner Press; 1990:291-314.
38. Cooper ML. Motivations for alcohol use among adolescents: Development and validation of a four-factor model. *Psychol Assessment* 1994;6(2):117-128. [doi: [10.1037/1040-3590.6.2.117](https://doi.org/10.1037/1040-3590.6.2.117)]
39. Kuntsche E, Knibbe R, Gmel G, Engels R. Replication and validation of the Drinking Motive Questionnaire Revised (DMQ-R, Cooper, 1994) among adolescents in Switzerland. *Eur Addict Res* 2006;12(3):161-168. [doi: [10.1159/000092118](https://doi.org/10.1159/000092118)] [Medline: [16778437](https://pubmed.ncbi.nlm.nih.gov/16778437/)]
40. Wurdak M, Dörfler T, Eberhard M, Wolstein J. Tagebuchstudie zu Trinkmotiven, Affektivität und Alkoholkonsum bei Jugendlichen. *Sucht* 2010;56(3-4):175-182. [doi: [10.1055/s-0031-1284504](https://doi.org/10.1055/s-0031-1284504)]
41. Stewart SH, Zack M. Development and psychometric evaluation of a three-dimensional Gambling Motives Questionnaire. *Addiction* 2008 Jul;103(7):1110-1117. [doi: [10.1111/j.1360-0443.2008.02235.x](https://doi.org/10.1111/j.1360-0443.2008.02235.x)] [Medline: [18554344](https://pubmed.ncbi.nlm.nih.gov/18554344/)]
42. Cooper ML, Shapiro CM, Powers AM. Motivations for sex and risky sexual behavior among adolescents and young adults: a functional perspective. *J Pers Soc Psychol* 1998 Dec;75(6):1528-1558. [Medline: [9914665](https://pubmed.ncbi.nlm.nih.gov/9914665/)]
43. Kuntsche E, le Mével L, Berson I. Development and Validation of the Motives for Listening to Music Questionnaire. *MLMQ* 2014 (forthcoming).
44. Kuntsche E, Kuntsche S. Development and validation of the Drinking Motive Questionnaire Revised Short Form (DMQ-R SF). *J Clin Child Adolesc Psychol* 2009 Nov;38(6):899-908. [doi: [10.1080/15374410903258967](https://doi.org/10.1080/15374410903258967)] [Medline: [20183672](https://pubmed.ncbi.nlm.nih.gov/20183672/)]
45. SosciSurvey. OnlineFragebogen URL: <https://www.soscisurvey.de/> [accessed 2014-07-28] [WebCite Cache ID 6ROmclFQq]
46. Riva G, Teruzzi T, Anolli L. The use of the internet in psychological research: comparison of online and offline questionnaires. *Cyberpsychol Behav* 2003 Feb;6(1):73-80. [doi: [10.1089/109493103321167983](https://doi.org/10.1089/109493103321167983)] [Medline: [12650565](https://pubmed.ncbi.nlm.nih.gov/12650565/)]
47. Thielsch MT. *Ästhetik von Websites. Wahrnehmung von Ästhetik und deren Beziehung zu Inhalt, Usability und Persönlichkeitsmerkmalen. Aesthetics of websites. Perception of aesthetics and its relationship to content, usability and personality trait*. Münster: Monsenstein und Vannerdat Wissenschaft; 2008.
48. Müller KW, Glaesmer H, Brähler E, Woelfling K, Beutel ME. Prevalence of internet addiction in the general population: results from a German population-based survey. *Behav Inform Technol* 2013 Aug 15;33(7):757-766. [doi: [10.1080/0144929X.2013.810778](https://doi.org/10.1080/0144929X.2013.810778)]
49. Wölfling K, Müller KW, Beutel M. [Reliability and validity of the Scale for the Assessment of Pathological Computer-Gaming (CSV-S)]. *Psychother Psychosom Med Psychol* 2011 May;61(5):216-224. [doi: [10.1055/s-0030-1263145](https://doi.org/10.1055/s-0030-1263145)] [Medline: [20878599](https://pubmed.ncbi.nlm.nih.gov/20878599/)]
50. Saß H, Wittchen HU, Zaudig M, Houben I. *Diagnostisches und Statistisches Manual Psychischer Störungen, Text revision (DSM-IV-TR)*. Göttingen: Hogrefe; 2003.
51. Dilling H, Mombour W, Schmidt MH. *International Classification of Diseases (ICD). German Version*. Bern: Huber; 2009.
52. Steffen S, Peukert P, Petersen KU, Batra A. Messverfahren zur Erfassung der Internetsucht. *Sucht* 2012 Dec;58(6):401-413. [doi: [10.1024/0939-5911.a000217](https://doi.org/10.1024/0939-5911.a000217)]
53. Bühner M. *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium; 2010.
54. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall; 2010.
55. Kline RB. *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press; 2005.
56. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. *Testing Structural Equation Models*. Newbury Park: Sage; 1993:136-161.
57. Nunnally JC, Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill; 1994.
58. Kuntsche E, Stewart SH, Cooper ML. How stable is the motive-alcohol use link? A cross-national validation of the Drinking Motives Questionnaire Revised among adolescents from Switzerland, Canada, and the United States. *J Stud Alcohol Drugs* 2008 May;69(3):388-396. [Medline: [18432381](https://pubmed.ncbi.nlm.nih.gov/18432381/)]
59. Cooper ML, Russell M, Skinner JB, Windle M. Development and validation of a three-dimensional measure of drinking motives. *Psychol Assessment* 1992;4(2):123-132. [doi: [10.1037/1040-3590.4.2.123](https://doi.org/10.1037/1040-3590.4.2.123)]

60. Kuntsche E, Knibbe R, Gmel G, Engels R. Why do young people drink? A review of drinking motives. *Clin Psychol Rev* 2005 Nov;25(7):841-861. [doi: [10.1016/j.cpr.2005.06.002](https://doi.org/10.1016/j.cpr.2005.06.002)] [Medline: [16095785](https://pubmed.ncbi.nlm.nih.gov/16095785/)]
61. Stumpp G, Stauber B, Reinl H. JuR - Einflussfaktoren, Motivation und Anreize zum Rauschtrinken bei Jugendlichen [Influencing factors, motivation and inducement for binge-drinking of adolescents]. Berlin: Bundesministerium für Gesundheit; 2009. URL: http://drogenbeauftragte.de/fileadmin/dateien-dba/DrogenundSucht/Alkohol/Downloads/Studie_Rauschtrinken_Uni_Tuebingen_090401_Drogenbeauftragte.pdf [accessed 2013-12-09] [WebCite Cache ID [6Lk1hcY0V](https://www.webcitation.org/6Lk1hcY0V)]

Abbreviations

ANOVA: analysis of variance

CFA: confirmatory factor analysis

CFI: comparative fit index

χ^2_{red} : reduced chi-square (chi-square divided by degrees of freedom)

CSV-R: screening scale for computer gaming behavior (“Screener zum Computerspielverhalten”)

DMQ-R: Drinking Motives Questionnaire Revised

DSM: Diagnostic and Statistical Manual of Mental Disorders

IAT: Internet Addiction Test

ICD: International Statistical Classification of Diseases and Related Health Problems

IMQ-A: Internet Motive Questionnaire for Adolescents

MOGQ: Motives for Online Gaming Questionnaire

OSVK-S: Screening measure for online addiction behavior in children and adolescents (“Online-Suchtverhalten-Skala”)

RMSEA: root mean square error of approximation

SRMR: standardized root mean squared residual

TLI: Tucker-Lewis index

Edited by G Eysenbach; submitted 28.03.14; peer-reviewed by Z Demetrovics, F Liga; comments to author 17.07.14; revised version received 09.08.14; accepted 16.08.14; published 09.10.14.

Please cite as:

Bischof-Kastner C, Kuntsche E, Wolstein J

Identifying Problematic Internet Users: Development and Validation of the Internet Motive Questionnaire for Adolescents (IMQ-A)

J Med Internet Res 2014;16(10):e230

URL: <http://www.jmir.org/2014/10/e230/>

doi: [10.2196/jmir.3398](https://doi.org/10.2196/jmir.3398)

PMID: [25299174](https://pubmed.ncbi.nlm.nih.gov/25299174/)

©Christina Bischof-Kastner, Emmanuel Kuntsche, Jörg Wolstein. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 09.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Comorbidity Structure of Psychological Disorders in the Online e-PASS Data as Predictors of Psychosocial Adjustment Measures: Psychological Distress, Adequate Social Support, Self-Confidence, Quality of Life, and Suicidal Ideation

Ali M AL-Asadi^{1,2}, MA, MSc; Britt Klein^{3,4,5}, DPsych (Clinical); Denny Meyer¹, DBL

¹School of Health Sciences, Swinburne University of Technology, Hawthorn, Australia

²Department of Arts and Education, Grande Prairie Regional College, Grande Prairie, AB, Canada

³DVC-Research & Innovation Portfolio; the School of Health Sciences and; the Collaborative Research Network, Federation University, Ballarat, Australia

⁴National Institute of Mental Health Research, The Australian National University, Canberra, Australia

⁵National eTherapy Centre, School of Health Sciences, Swinburne University of Technology, Hawthorn, Australia

Corresponding Author:

Ali M AL-Asadi, MA, MSc

Department of Arts and Education

Grande Prairie Regional College

10726 - 106 Ave.,

Grande Prairie, AB,

Canada

Phone: 1 780 539 2061

Fax: 1 780 538 0960

Email: aalasaki@gprc.ab.ca

Abstract

Background: A relative newcomer to the field of psychology, e-mental health has been gaining momentum and has been given considerable research attention. Although several aspects of e-mental health have been studied, 1 aspect has yet to receive attention: the structure of comorbidity of psychological disorders and their relationships with measures of psychosocial adjustment including suicidal ideation in online samples.

Objective: This exploratory study attempted to identify the structure of comorbidity of 21 psychological disorders assessed by an automated online electronic psychological assessment screening system (e-PASS). The resulting comorbidity factor scores were then used to assess the association between comorbidity factor scores and measures of psychosocial adjustments (ie, psychological distress, suicidal ideation, adequate social support, self-confidence in dealing with mental health issues, and quality of life).

Methods: A total of 13,414 participants were assessed using a complex online algorithm that resulted in primary and secondary Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision) diagnoses for 21 psychological disorders on dimensional severity scales. The scores on these severity scales were used in a principal component analysis (PCA) and the resulting comorbidity factor scores were related to 4 measures of psychosocial adjustments.

Results: A PCA based on 17 of the 21 psychological disorders resulted in a 4-factor model of comorbidity: anxiety-depression consisting of all anxiety disorders, major depressive episode (MDE), and insomnia; substance abuse consisting of alcohol and drug abuse and dependency; body image-eating consisting of eating disorders, body dysmorphic disorder, and obsessive-compulsive disorders; depression-sleep problems consisting of MDE, insomnia, and hypersomnia. All comorbidity factor scores were significantly associated with psychosocial measures of adjustment ($P < .001$). They were positively related to psychological distress and suicidal ideation, but negatively related to adequate social support, self-confidence, and quality of life.

Conclusions: This exploratory study identified 4 comorbidity factors in the e-PASS data and these factor scores significantly predicted 5 psychosocial adjustment measures.

Trial Registration: Australian and New Zealand Clinical Trials Registry ACTRN121611000704998; http://www.anzctr.org.au/trial_view.aspx?ID=336143 (Archived by WebCite at <http://www.webcitation.org/618r3wvOG>).

KEYWORDS

comorbidity; comorbidity factors, comorbidity dimensions, structure of comorbidity; psychosocial adjustment; disorders; suicidal ideation; quality of life (QoL); self-confidence; distress; social support; online therapy; e-mental health; generalized anxiety disorder; obsessive-compulsive disorder; social anxiety disorder; posttraumatic stress disorder; PD; major depressive episode; anxiety disorders; insomnia, hypersomnia

Introduction

Background

The co-occurrence of multiple psychological disorders (comorbidity) is a common and serious problem that needs to be better researched and understood. Kessler et al [1] found the lifetime prevalence of any disorder to be 46.4%, whereas 27.7% and 17.3% were the lifetime prevalences of 2 or more and 3 or more disorders, respectively. Kessler et al [2] found the 12-month prevalence of any disorder to be 26.2%, whereas 5.8% met the criteria for 2 disorders and 6.0% met the criteria for more than 2 disorders. They also concluded that more than 40% of those who received 1 diagnosis also met the diagnostic criteria for a second diagnosis over a 12-month period [2]. Naturally, those suffering from multiple psychological disorders would more likely require greater assistance from their therapists should they seek help. Indeed, several studies concluded that comorbidity was consistently related to poor prognosis, higher rate of suicide, and high demand for professional help [3,4]. In addition, comorbidity has also been shown to be strongly related to symptom severity [2].

Although comorbidity has been studied extensively over the past 30 years, most studies on the structure of comorbidity have used 6-12 disorders at most and recommended that future research should include more disorders [5,6]. Furthermore, the use of dimensional rather than a discrete or dichotomous approach was suggested to solve the problems of generalizability and detection of the correlates of more severe mental disorders [7]. In addition, avoiding the reliance on threshold diagnoses in favor of consideration of subthreshold information has been recommended [8].

Krueger [7] argued for using a dimensional rather than a discrete/dichotomous approach and for including patients who met the criteria for multiple disorders when studying the structure of comorbidity because such patients were typical and more severely impaired. Consequently, Krueger argued that basing research designs on such samples would increase generalizability and the detection of correlates of more severe mental disorders. To this end, Krueger performed factor analysis on 10 mental disorders based on the *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition, Revised; *DSM-III-R*) using the National Comorbidity Survey (NCS) sample of 8098 (collected between 1990 and 1992) and found 2 factors: an externalizing factor which included alcohol and drug dependency and antisocial personality disorder and an internalizing factor which was broken down into an anxious-misery factor, including depression, dysthymia, generalized anxiety disorder (GAD); and a fear factor, including social anxiety disorder (SAD), simple/specific phobia, agoraphobia, and panic disorder (PD). This model was criticized

for the restricted range of 10 disorders, the reliance on lifetime diagnoses, and the exclusive reliance on threshold diagnoses without consideration of the subthreshold diagnostic information [8]. However, the structure of Krueger's factors was later supported by several studies [2,6,9-12].

To review the literature on the comorbidity and the structure of comorbidity of all 21 disorders will require a great deal of space and is beyond the scope of this report. Therefore, we shall limit this summary review to disorders that have received the most attention and are the most prevalent: anxiety disorders, major depressive disorder (MDD), eating disorders, and alcohol and drug abuse and dependency disorders.

Anxiety and Major Depressive Disorders

Comorbidity of anxiety disorders and MDD is among the most prevalent comorbid psychiatric conditions [13,14]. Estimates of comorbid anxiety disorders and MDD range from 15.9% to 61.9% in children and adolescents [15] and from 14.5% [16] to 57% [17-20] in specific populations of adults. In addition, sleep problems such as insomnia and hypersomnia have been consistently found to be associated with anxiety and mood disorders [21-25]. These studies suggest the existence of an anxiety-depression-insomnia comorbidity factor.

The co-occurrence of MDD and anxiety disorders has many clinical implications. Several studies found outpatients with MDD and at least 1 co-occurring anxiety disorder were significantly more likely to suffer from insomnia [21], personality disorders [26], more severe and chronic depression [16,27], and engage in suicidal ideation [18] and suicidal behavior [28] than individuals with only a MDD diagnosis.

GAD is highly correlated with mood disorders and has slightly lower correlations with anxiety disorders [29]. The tendency for GAD to be highly comorbid with other anxiety and depressive disorders has led some researchers to question its discriminant validity among diagnostic categories [30]. Researchers have pointed out that some of the core features of GAD are usually present in varying degrees in all anxiety disorders as well as mood disorders, detracting from the overall discriminant validity and reliability of GAD as a principal diagnosis. Factor analysis has also revealed a closer link between GAD and depressive disorders such as major depression, dysthymia, and major depressive episode (MDE) (the internalizing anxious-misery dimension) than between GAD and anxiety disorders such as social phobia, simple/specific phobia, PD, agoraphobia, and obsessive-compulsive disorder (OCD)—the internalizing fear dimension [5,6]. Some researchers suggested that GAD would be better categorized with the mood disorders whereas others suggested a complete revision of the hierarchical structure of the *DSM-IV* in which all anxiety and mood disorders would be grouped together and

then partitioned into 3 subclasses with MDD, GAD, dysthymic disorder, and posttraumatic stress disorder (PTSD) forming a subclass of distress disorders [31-33].

Conversely, a more recent longitudinal study supported the *DSM-IV-TR* classification of GAD as an anxiety disorder. In an effort to determine whether GAD would be better classified as a mood disorder rather than an anxiety disorder, Beesdo et al [34] examined the risk patterns, incidence, developmental features, and comorbidity of anxiety and depressive disorders in 3021 individuals in a prospective longitudinal study conducted over a period of more than 10 years. They concluded that GAD was linked more closely to anxiety disorders than mood disorders in terms of risk associations in familial, temperamental, personality, and environmental variables. Moreover, temporal comorbidity of GAD showed a strong association between GAD and other anxiety disorders.

One of the earliest reviews of the literature on comorbidity among anxiety disorders is the one conducted by Brown and Barlow [30]. They highlighted several points: diagnoses of PD with or without agoraphobia and GAD were associated with some of the highest comorbidity rates between psychiatric conditions, core features of PD with or without agoraphobia and GAD were shared to some extent with all anxiety disorders, and substance abuse followed by GAD were the most commonly comorbid lifetime and current disorders experienced by war veterans with or without PTSD. They also concluded that, as with GAD, the discriminant validity of PD with or without agoraphobia was questionable because some of the central features in PD with or without agoraphobia (eg, anxious apprehension) were present in varying degrees in all anxiety and mood disorders. Subsequent studies found essentially the same associations and comorbidities—PD frequently co-occurs with other anxiety disorders [35-37].

Depression, Anxiety, and Eating Disorders

Depressive and anxiety disorders are also frequently reported in those diagnosed with eating disorders. The lifetime prevalence of major depression ranges from 50% to 71% in anorexia nervosa and 50% to 65% in bulimia nervosa [38-41]. When the age was restricted to 12-18 years, 60% of adolescent girls with anorexia nervosa reported comorbid mood disorder [42]. On the other hand, when the age was broadened to include persons aged 11-68 years, 92% of the large female sample with anorexia nervosa had unipolar depression [43].

Although studies investigating the relationship between eating and anxiety disorders have produced somewhat mixed results [44-46], several researchers have found high rates of comorbidity between eating and anxiety disorders [43,47-50]. The few studies that employed control groups found significant comorbidity between anxiety disorders and anorexia nervosa and bulimia nervosa populations in comparison to non-eating disordered controls [46,51-54].

Moreover, Allen and Hollander [55] and Cororve and Gleaves [56] pointed out the significant overlap between the features of eating disorders (particularly anorexia nervosa) and body dysmorphic disorder (BDD). In both cases, body image distortion, excessive concerns and unreasonable preoccupation

about physical appearance, and stressful dissatisfaction with one's body are common features. Moreover, similar to the behaviors of individuals with OCD, excessive concerns and preoccupation over one's body lead to anxiety-provoking obsessive thoughts about one's self-image which might force the individual to engage in ritualistic-like behaviors to reduce the generated anxiety [55,57,58]. Therefore, all these studies suggest the existence of an eating-anxiety comorbidity factor and/or a body image-eating disorder comorbidity factor.

Anxiety and Substance Abuse and Dependency Disorders

The relationship between psychological disorders and the use of drugs and alcohol has long been established [59-61]. McGovern et al [62] found substance use disorders were present in 42%, 27%, and 20% of patients diagnosed with mood disorders, anxiety disorders, and antisocial personality disorder, respectively. Krueger [7] found alcohol and drug dependence and antisocial personality disorder to make up the externalizing factor. Katz et al [63] performed confirmatory factor analysis on the same NCS sample used by Krueger [7], but this time with the inclusion of alcohol and drug abuse and dependence. They found alcohol and drug abuse to have strong negative loadings, whereas alcohol and drug dependency were found to have strong positive loadings on the externalizing factor. In addition, they found alcohol and drug abuse to have substantial loadings on the anxious-misery subfactor of the internalizing factor. They concluded that there appears to be a group of individuals with alcohol and drug abuse that is distinct in terms of comorbidity patterns and etiology from those with alcohol and drug dependency and from those with anxious-misery disorders (MDE, dysthymia, GAD). There appears to be some support, at least as far as alcohol is concerned, for the distinction between alcohol abusers and alcohol dependents in that, unlike abusers, alcohol dependency consists of a single latent dimension whereas alcohol abuse consists of many underlying latent dimensions [64]. These studies suggest the presence of a substance dependence comorbidity factor.

In summary, it appears that comorbidity of psychological disorders is a prevalent phenomenon with almost half of those receiving a diagnosis also receiving multiple diagnoses. Comorbidity is associated with more demands for services, poorer prognosis, higher rate of suicide, greater levels of severity, and poorer adjustment. Most studies on the structure of comorbidity focused on a few psychological disorders. Anxiety disorders, MDD, and eating disorders are the most widely researched. Results show that comorbidity of anxiety and MDDs are most prevalent. MDD was comorbid with GAD, OCD, PTSD, PD with or without agoraphobia, SAD, simple/specific phobia, insomnia, hypersomnia, and alcohol and drug abuse disorders. Anxiety disorders are particularly comorbid with GAD as well as with one another. In addition, significant comorbidity has been found between anxiety, depression, and eating disorders. Finally, although alcohol and drug abuse disorders appear to be comorbid with MDE and some anxiety disorders, alcohol and drug dependency disorders form a separate group with antisocial personality disorder.

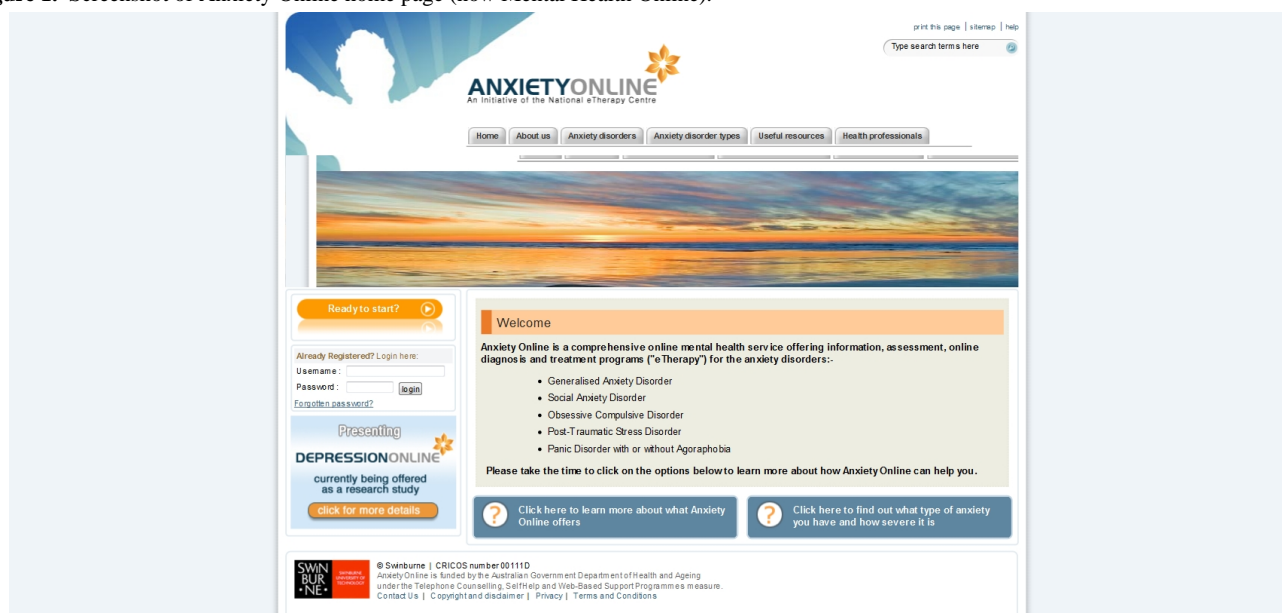
The Mental Health Online (previously Anxiety Online) electronic psychological assessment screening system (e-PASS; an online psychological assessment screening system assessing for 21 *DSM-IV-TR* disorders) dataset (see [65,66] for further information), contains information on more than 13,000 individuals who have received 21 single and/or multiple psychological diagnoses based on dimensional measures referred to as clinical disorder severity scales. These elements of the e-PASS dataset provide a unique opportunity to examine the structure of comorbidity of 21 psychological disorders based on dimensional scales that reflect information that covers subthreshold to postthreshold diagnoses.

The purpose of this exploratory study was 2-fold. Firstly, we used the dimensional severity scales of the e-PASS data to explore the underlying structural dimensions of comorbidity (henceforth, comorbidity factors) of the 21 psychological

disorders diagnosed by the e-PASS part of the Mental Health Online platform (Figure 1) [67]. Secondly, we validated the resulting comorbidity factors by determining their relative performance for predicting psychological distress as measured by Kessler-6 total score, suicidal ideation, quality of life, level of self-confidence, and adequate social support. We predicted that comorbidity factor scores would be positively related to suicidal ideation and Kessler-6 total score, but negatively related to quality of life, self-confidence, and adequate social support.

For those diagnosed with more than 1 disorder, it has been found that the number of diagnosed psychological disorders is positively related to the level of psychological distress [68] and to suicidal ideation [3,4], whereas negatively related to patients' level of self-confidence, quality of life [69,70], and social support [71,72]. Therefore, any comorbidity factor scores should adequately predict these validation variables.

Figure 1. Screenshot of Anxiety Online home page (now Mental Health Online).



Methods

Procedure

The Mental Health Online platform consists of 4 centers: psychoeducational information, psychological screening and assessment, online treatment, and health care practitioner training. The psychoeducational center is a website that provides psychoeducational information. The assessment center contains the e-PASS. The treatment center provides and manages the 5 anxiety treatment programs and the more recent depression, bulimia nervosa, and insomnia treatment programs. The training center provides the eTherapist training programs and the health care practitioner portal. The online psychological assessment and referral program, e-PASS, includes a variety of demographic and personal questions, the Kessler-6 [73], and the Suicidal Warnings (ideation) measure, as well as the online diagnostic program. Individuals can access the Mental Health Online service from anywhere in the world provided they have an Internet connection. People can complete e-PASS if they are interested in the psychological assessment function and/or if

they are interested in online treatment. Those who want to undertake the e-PASS are first required to register and consent to the Mental Health Online terms and conditions [67]. The procedures for collecting and reporting the Anxiety Online data were approved by the Swinburne University Human Research Ethics Committee.

Diagnostic Assessment

Based on an individual's response to some of the e-PASS questions, a person may be given a primary diagnosis and/or multiple secondary diagnoses. Primary or secondary diagnosis is determined by the reported presence of *DSM-IV-TR* symptoms and the average score on severity scales, each of which assesses the level of distress and interference caused by the symptoms of a particular disorder. A total of 21 clinical disorders are assessed by the e-PASS (see [65,66] for more details).

The disorder-specific severity score is the average of the scores on 6 questions that assess how distressed and how many of the symptoms of a given disorder interfere in one's life. For example, an individual who meets the diagnostic symptomatology of PD has to rate each of the following

questions on a scale from 0 indicating no distress or interference to 8 indicating extremely severe distress or interference as indicated subsequently:

Currently, how bothered or distressed are you by your panic attack(s) and or worries about having another panic attack, and or any changes in your behavior/routine because of these panic attacks, in each of the following areas of your life:

1. Everyday living (eg, shopping, cleaning the house, preparing food)
2. Social (eg, meeting friends or family, forming/maintaining relationships)
3. Work or study (eg, attending work or classes, performing tasks)

Currently, how much interference do your panic attack(s), and or worries about having another panic attack, and or changes in your behavior/routine because of these panic attacks cause, in each of the following areas of your life:

1. Everyday living (eg, shopping, cleaning the house, preparing food)
2. Social (eg, meeting friends or family, forming/maintaining relationships)
3. Work or study (eg, attending work or classes, performing tasks)

A person who does not endorse the initial *DSM-IV-TR* symptom criteria questions for a particular disorder is not presented with the questions assessing their level of distress and interference of those symptoms and is assigned a severity score of zero. Those who do endorse the initial *DSM-IV* symptom criteria questions for a particular disorder are presented with the 6 distress and interference questions allowing the calculation of a mean severity score ranging from 0 to 8. A mean distress and interference severity score of 3.5 or greater is considered sufficient to warrant a clinical diagnosis. Those whose mean distress and interference severity scores are less than 3.5 are considered to warrant a subclinical (or subthreshold) diagnosis.

The e-PASS diagnostic system was modeled after and informed by the Anxiety Disorders Interview Schedule (ADIS) clinician rating scale (Likert scale: 0=no symptoms, 4=mild presence of the disorder, 8=very severe presence of the disorder). Most "total scores" would not be a whole number because the system used 6 rating scales and then averaged them. Consequently, the final scores were rounded up or down to the nearest whole number. Thus, a 4 is the typical score by a clinician using the ADIS that indicates the "presence" of a disorder. However, considering the decimal places resulting from the e-PASS averaging of the 6 rating scales, those scoring 3.5 or more were deemed clinical.

The psychometric properties of the e-PASS measures were shown to have high test-retest reliability and reasonable convergent validity with the structured clinical interviews (D Nguyen, unpublished PhD thesis, Swinburne University, 2013). However, the small sample size and some disagreement with the structured clinical interviews in terms of the severity levels required for a clinical diagnosis suggest that further validation studies with large sample sizes are needed.

For the purpose of this work, we did not use the *DSM-IV-TR* criteria for clinical diagnosis. Instead, we used the previously described severity score without consideration of the level of severity that would warrant clinical diagnosis. The average severity scores provide the opportunity to examine the structure of comorbidity of 21 psychological disorders based on a dimensional scale that reflects information that covers subthreshold to postthreshold diagnoses. This dimensional approach should solve the problems of generalizability and detection of the correlates of more severe mental disorders as suggested by Krueger [7]. Also, the inclusion of subthreshold symptomatology should avoid the risk of finding artifactual explanations as was suggested by Wittchen et al [8].

Participants

A total of 13,414 individuals completed the e-PASS phase between October 2009 and October 2012 and received at least 1 clinical diagnosis. The sample consisted of 3974 (29.63%) males whose age ranged between 18-85 years with a mean of 36.88 (SD 12.59) years, and 9440 (70.37%) females whose ages ranged between 18-86 years with a mean of 33.66 (SD 11.57) years. A total of 749 (5.58%) participants received a clinical diagnosis (severity score greater than 3.5) of only 1 disorder, leaving 12,665 (94.42%) participants who were classified as having a clinical or subclinical diagnosis for 2 or more of the 21 disorders assessed by e-PASS.

In this study, the score that each participant received before treatment on the following variables were used to validate the resulting comorbidity factors: Kessler-6 total score, suicidal ideation, adequate social support ("Do you feel you have adequate level of social support or engagement in social and/or community activities?"), self-confidence ("Please rate your overall level of self-confidence when it comes to managing your own mental health"), and quality of life ("Please rate your overall quality of life"). The suicidal ideation and adequate social support measures were based on binary yes/no answer responses. The level of self-confidence and the quality of life measures were based on a Likert scale from "very poor" to "very good." For the purpose of this study, the last 2 measures were converted to binary scale, with all those who gave a rating of very poor and poor coded as zero and those who gave a rating for good and very good coded as "1".

Analysis

A series of principal component analyses were conducted to develop new comorbidity factors from the severity scores described previously. The descriptive statistics are reported for the 21 severity scores and for the validation variables used to validate the new comorbidity factors.

The pervasive nature of depression and anxiety disorders meant that a typical factor analysis was not appropriate for the extraction of the comorbidity factors. In particular, it was not appropriate to search for a simple structure for the loadings of the severity scores on the final comorbidity factors. Instead, the goal was to derive the underlying dimensions that explained as much of the variation in the severity scores as possible while recognizing the importance of depression and anxiety in the majority of the comorbidity factors. The relative accuracy of

the severity score measures suggested that a principal component analysis (PCA) should be used to construct the comorbidity factors, rather than factor analysis with principal axis factoring or maximum likelihood extraction of factors. This choice was supported when comparative analyses performed with maximum likelihood and principal axis factoring extraction methods failed to produce results supported in the literature.

Monte Carlo simulation (parallel analysis) [74] was carried out to determine the optimum number of components at each step, based on the work of O'Connor [75]. Initially an oblimin rotation was applied to the factors, allowing correlation between the resulting comorbidity factors, but when these correlations were found to be very weak, the oblimin rotation was replaced by a varimax rotation ensuring that the resulting comorbidity factors were independent of each other. At each step of the analysis, disorders that loaded weakly on all factors were removed.

In the second part of the analysis, comorbidity factor scores were validated using linear and logistic multiple regression to

test their relationships with psychological distress (as measured by the Kessler-6), suicidal ideation, adequate social support, self-confidence, and quality of life. All analyses were carried out using SPSS version 20 (IBM Corp, Armonk, NY, USA).

Results

Principal Component Analysis

The number of clinical diagnoses among those who received more than 1 diagnosis ranged from 1 to 16 with a mean of 5.07 (SD 2.41). The number of diagnoses was significantly and positively correlated with the psychological distress Kessler-6 total score ($r=.54$, $P<.001$), negatively with quality of life ($r=-.32$, $P<.001$), negatively with self-confidence ($r=-.29$, $P<.001$), and negatively with level of adequate social support ($r=-.24$, $P<.001$). The 21 disorders and all descriptive statistics are shown in Table 1. The validation variables and their descriptive statistics are shown in Table 2.

Table 1. The frequency of severity of 21 psychological disorders.

e-PASS Assessment	Severity of disorder (N=13,414)		Mean (SD)
	Clinical	Subclinical	
Psychological disorders, n (%)			
Panic disorder (PD) with or without agoraphobia	2970 (22.14)	1809 (13.49)	1.43 (2.21)
Agoraphobia without a history of PD	1107 (8.25)	1447 (10.79)	0.61 (1.47)
Social anxiety disorder (SAD)	3504 (26.12)	4031 (30.05)	1.92 (2.20)
Specific phobia	1823 (13.59)	3055 (22.77)	1.10 (1.86)
Generalized anxiety disorder (GAD)	4464 (33.28)	4366 (32.55)	2.35 (2.28)
Obsessive-compulsive disorder (OCD)	1033 (7.70)	2730 (20.35)	0.74 (1.52)
Posttraumatic stress disorder (PTSD)	1461 (10.89)	2974 (22.17)	0.96 (1.74)
Major depressive episode (MDE)	5131 (38.25)	4188 (31.22)	2.63 (2.34)
Anorexia nervosa	41 (0.31)	64 (0.48)	0.02 (0.31)
Bulimia nervosa	804 (5.99)	931 (6.94)	0.46 (1.46)
Binge eating disorder	316 (2.36)	682 (5.08)	0.21 (0.89)
Substance dependence			
Cannabis	127 (0.95)	839 (6.25)	0.14 (0.63)
Stimulants	48 (0.36)	492 (3.67)	0.07 (0.44)
Opioids	29 (0.22)	313 (2.33)	0.05 (0.34)
Sedatives	104 (0.78)	1218 (9.08)	0.18 (0.65)
Alcohol dependence	321 (2.39)	2201 (16.41)	0.38 (0.96)
Somatization disorder	179 (1.33)	51 (0.38)	0.09 (0.70)
Body dysmorphic disorder	1503 (11.20)	649 (4.84)	0.71 (1.79)
Problem/pathological gambling	49 (0.37)	457 (3.41)	0.07 (0.42)
Insomnia	4049 (30.18)	4546 (33.89)	2.20 (2.23)
Hypersomnia	1060 (7.90)	827 (6.17)	0.53 (1.50)

Table 2. The frequency of and descriptive statistics of the validation variables.

Validation variables	Frequency (N=13,414)	Mean (SD)
Suicidal ideation (yes), n (%)	5079 (3.86)	
Quality of life (high), n (%)	6877 (51.27)	
Adequate social support (high), n (%)	6024 (44.91)	
Self-confidence (high), n (%)	3888 (28.98)	
Pre-Kessler-6 (total score)		17.49 (4.90)

All but 6 of 210 severity score correlations among disorder pairs were positive, and all but 11 of 210 correlations were statistically significant ($P < .05$). The Kaiser-Meyer-Olkin (KMO) of 0.80 confirmed that there was sufficient correlation in the data to warrant a PCA analysis of this nature.

To determine the optimum number of comorbidity factors, both methods of Monte Carlo simulation, random data and permutation, indicated the presence of 4 underlying dimensions, as suggested by the literature review. Disorders that loaded weakly ($< .3$) on all factors were removed which resulted in the removal of problem/pathological gambling, somatization disorder, and binge eating disorder because none of these disorders featured prominently in the comorbidity literature. The resulting PCA model accounted for 42.9% of the total variance in the remaining 18 severity scores when an oblimin rotation was used. However, the component correlation matrix showed all correlations to be less than .3 in absolute value, suggesting that a simpler solution with 4 distinct comorbidity factors was possible with a varimax rotation.

The penultimate 4-component model indicated that the fourth component consisted of a negative loading of PD with or without agoraphobia ($-.59$) and a positive loading of agoraphobia without a history of PD (.86). This was a reflection of the mutual exclusivity of these 2 classifications. That is, according to *DSM-IV-TR*, a person can only be diagnosed with 1 or the other but not both of these disorders at the same time. Also, although a KMO of 0.8 was indicative of the suitability of the dataset for

factor analysis, examination of the anti-image of the correlation matrix revealed that the measures of sampling adequacy for each individual variable ranged from 0.69 to 0.91 except for the agoraphobia without a history of PD, which was rather poor at 0.4 suggesting that this variable should be removed. For these reasons, it was decided to remove agoraphobia without a history of PD from the analysis and develop a new model.

After removing agoraphobia without a history of PD, the KMO value increased to 0.85 and a new 4-component model emerged accounting for 43.9% of the variance in severity scores. As shown in [Table 3](#), the final 4-component model based on a varimax rotation provided support for the 4 comorbidity factors suggested by the literature. The first component included anxiety disorders (specific phobia, PD with or without agoraphobia, GAD, SAD, PTSD, and OCD) with MDE and insomnia having the highest loading on this anxiety-depression factor. The second component showed moderate loadings for alcohol and other drug dependency-related disorders (stimulant dependence, cannabis dependence, opioid dependence, alcohol dependence, and sedative dependence) confirming the existence of a substance abuse comorbidity factor. The third component exhibited high to moderate positive loadings for bulimia nervosa, BDD, anorexia nervosa, and OCD, confirming the presence of a body image-eating factor. The fourth component showed high to moderate positive loadings for MDE, insomnia, and hypersomnia, confirming the existence of a depression-insomnia comorbidity factor.

Table 3. Component weights for principal component analysis with varimax rotation.

Disorder severity	Component			
	Anxiety-depression	Substance abuse	Body image-eating	Depression-sleep problems
Specific phobia	.715			
PD with or without agoraphobia	.695			
GAD	.654			
SAD	.622			
MDE	.541			.509
PTSD	.537			
Insomnia	.481			.454
OCD	.424		.356	
Stimulant dependence		.638		
Cannabis dependence		.596		
Opioid dependence		.513		
Sedative dependence		.499		
Alcohol dependence		.493		
Bulimia nervosa			.791	
BDD			.705	
Anorexia nervosa			.408	
Hypersomnia				.745

Comorbidity Factor Scores

Principal component scores representing these 4 comorbidity factors were calculated for each participant. The first component represents the anxiety-depression comorbidity factor. The second component represents the substance abuse comorbidity factor. The third component represents the body image-eating comorbidity factor. The fourth component represents the depression-sleep problems comorbidity factor.

The 4 comorbidity factor scores were then tested for their relationships with the validation measures. These uncorrelated factor scores served as the independent variables for multiple regression analysis with the Kessler-6 total score serving as the dependent variable. The results are shown in Table 4. The regression coefficients for the 4 comorbidity scores were positive and statistically significant ($P<.001$), explaining 44% of the variation in the Kessler-6 total score.

Table 4. Results of regression analysis for the comorbidity factor scores and Kessler-6 total score

Comorbidity factors	B	SE	t_1	P	R	R^2
					.66	.44
Anxiety-depression	2.49	.03	78.4	.001		
Substance abuse	0.43	.03	13.6	.001		
Body image-eating	0.97	.03	30.5	.001		
Depression-sleep problems	1.81	.03	57.1	.001		

In addition, these comorbidity factor scores were related to 4 other variables: suicidal ideation, adequate social support, self-confidence, and quality of life using binary logistic regression as shown in Table 5. The final binary logistic regressions with forward selection resulted in significant odds ratios for all 4 comorbidity factor scores. On average, increased odds for experiencing suicidal ideation of 108%, 36%, 48%,

and 91%; for inadequate social support of 71%, 10%, 16%, and 49%; for having low self-confidence of 30%, 13%, 52%, and 64%; and for poor quality of life of 31%, 22%, 33%, and 81% for each additional point scored on anxiety-depression, substance abuse, body image-eating, and depression-sleep problems factors, respectively.

Table 5. Binary logistic regression for suicidal ideation, inadequate social support, low self-confidence, and poor quality of life in relation to comorbidity factor scores.

Comorbidity factors	Wald (<i>df</i> =1)	<i>P</i>	OR (95% CI)
Suicidal ideation			
Anxiety-depression	1139.11	.001	2.08 (1.99-2.17)
Substance abuse	177.48	.001	1.36 (1.30-1.42)
Body image-eating	338.30	.001	1.48 (1.42-1.54)
Depression-sleep problems	886.34	.001	1.91 (1.83-2.00)
Constant	688.91	.001	0.59
Inadequate social support			
Anxiety-depression	650.90	.001	1.71 (1.64-1.78)
Substance abuse	21.65	.001	1.10 (1.06-1.14)
Body image-eating	55.50	.001	1.16 (1.11-1.20)
Depression-sleep problems	377.74	.001	1.49 (1.43-1.55)
Constant	183.94	.001	1.28
Low self-confidence			
Anxiety-depression	805.52	.001	2.30 (2.17-2.44)
Substance abuse	23.41	.001	1.13 (1.08-1.19)
Body image-eating	198.49	.001	1.52 (1.43-1.61)
Depression-sleep problems	335.03	.001	1.64 (1.56-1.73)
Constant	2287.57	.001	3.07
Poor quality of life			
Anxiety-depression	1267.24	.001	2.31 (2.20-2.42)
Substance abuse	77.36	.001	1.22 (1.17-1.27)
Body image-eating	174.99	.001	1.33 (1.27-1.38)
Depression-sleep problems	710.34	.001	1.81 (1.74-1.89)
Constant	0.66	.42	1.02

These results confirm that the 4 comorbidity factor scores are associated with the validation variables that are commonly associated with psychological distress, thereby providing some validation for these underlying dimensions of comorbidity.

Discussion

Initial Structure

Using the scores on dimensional scales measuring the severity of 21 psychological disorders, the underlying structure of comorbidity of these disorders was examined using the method of PCA. Somatization disorder, pathological gambling, and binge eating disorder were systematically removed from analysis because of weak loadings of less than 0.3 on all components. The penultimate model consisted of 4 components, the last of which consisted of only a negative loading for PD with or without agoraphobia and a positive loading for agoraphobia without a history of PD, reflecting the mutual exclusivity of these 2 diagnoses. According to *DSM-IV-TR*, a person can only be diagnosed with 1 or the other but not both disorders at the same time. This also suggests that there are differences between the 2 groups even though the 2 groups share common etiological

factors. This finding is consistent with the finding of Katz et al [63], but is contrary to the new *DSM-5* criterion that no longer allows agoraphobia without PD. However, Katz et al [63] found a negative relationship between PD with or without agoraphobia and agoraphobia without a history of PD, but Krueger [7] found a positive relationship between these 2 disorders. In view of these previous findings and the high component loadings for only these 2 disorders on our fourth component, we removed agoraphobia without a history of PD producing a new factor model containing 17 disorders loading on 4 reconfigured components.

Comorbidity Structure

The underlying dimensional structure of the 17 psychological disorders consisted of 4 factors that accounted for 43.9% of the total variance. The first comorbidity factor, labeled anxiety-depression, consisted of all anxiety disorders (specific phobia, PD with or without agoraphobia, GAD, SAD, PTSD, and OCD) with MDE and insomnia. The second comorbidity factor, labeled substance abuse, consisted of alcohol and substance dependency disorders (alcohol dependence, cannabis dependence, stimulant dependence, opioid dependence, and

sedative dependence). The third comorbidity factor, labeled body image–eating, consisted of body image and eating problems (bulimia nervosa, BDD, anorexia nervosa, and OCD). The fourth comorbidity factor, labeled depression–sleep problems consisted of MDE, insomnia, and hypersomnia. It is important to note that these comorbidity factors are unlikely to be a reflection of the inability of these dimensional severity scales to adequately distinguish between the different disorders. It is the lack of correlation between these 4 factors that suggests that these clusters of disorders are distinct rather than the severity scales lacking the specificity to distinguish between these disorders.

We should also emphasize here that we are not aware of any work on comorbidity or the structure of comorbidity of psychological disorders in the context of online programs. Therefore, this work may be the first in addressing this area. Consequently, we are comparing the results of this study with results based on using in-clinic samples. Although there is no logical reason to expect the comorbidity of psychological disorders to be different for an online population from an in-clinic population, caution is warranted.

These results are generally supportive of Krueger’s model [7]. The substance abuse factor corresponds with Krueger’s externalizing factor minus the antisocial personality disorder (not included in our data). The fact that all alcohol and drug dependency classifications are highly loaded on 1 factor implies that they may be largely determined by a single latent dimension. This finding is consistent with the work of Katz et al [63] who suggested that alcohol and drug dependency have a common latent dimension and the work of Slade et al [64] who concluded that alcohol dependency consisted of 1 latent dimension.

The anxiety-depression factor corresponds with Krueger’s internalizing factor [7]. However, our results do not support the separation of the internalizing factor into anxious-misery (depression, dysthymia, GAD) and fear (SAD, specific phobia, agoraphobia, PD) subfactors [7]. This could be due to the differences between our dataset and Krueger’s dataset [7]. Our results are based on 17 disordered classifications that were assessed on dimensional scales that included subthreshold scores whereas Krueger’s data [7] were based on 10 disorders that were assessed on scales that relied exclusively on threshold diagnoses.

It has long been known that anxiety disorders tend to be comorbid with depression as well as each other. Our results are consistent with this notion that anxiety disorders and depression share many common symptoms. The inclusion of insomnia in this group of disorders is understandable because it is often the case that people who are suffering from insomnia are also suffering from 1 or more anxiety and/or mood disorders [21,22]. This assertion becomes even clearer when we consider the depression-sleep problem factor. Insomnia and hypersomnia are sometimes associated with increases in anxiety and depression [23-25]. Although insomnia or hypersomnia could be symptoms of anxiety or mood disorders, they could also lead to the experience of anxiety and/or depressive symptoms.

The body image-eating factor consists of disorders that are related to the perception of body image. People suffering from

these disorders are not satisfied with the way they look, but each group deals with this dissatisfaction differently. The inclusion of OCD with these disorders is understandable and was predicted. Sufferers of eating disorders and body dysmorphic disorder share excessive concerns and preoccupation about physical appearance coupled with dissatisfaction with their bodies that commonly lead them to develop a distorted self-image. The anxiety-provoking obsessive thoughts about their appearance force them to engage in ritualistic behaviors to reduce their anxiety. This pattern of anxiety-provoking obsessive thoughts reduced by ritualistic-like behaviors is typical of OCD sufferers. Those with eating disorders obsess over their body weight and engage in behaviors that reduce their weight, such as limiting food intake, vigorous exercises, and using laxatives and/or forced purging. Those with BDD obsess over their appearance and engage in behaviors that reassure them, such as frequently checking with others in mirrors, comparing themselves to others, camouflaging themselves, and seeking cosmetic surgeries. These results are consistent with findings of Allen and Hollander [55] and Cororve and Gleaves [56], and with the finding of Rauch et al [57] who indicated that the same brain structures were implicated in BDD and OCD, and Phillips [58] who suggested that treatment for OCD is useful for BDD. However, it should be noted that, aside from OCD, our results do not support the existence of an eating-anxiety comorbidity factor.

Finally, now that the *DSM-5* has been introduced, there will be a need to revise the e-PASS assessment tool based on the new criteria. This revision will likely be needed for all diagnostic measures and for other previous comorbidity research. However, for some disorders, such as MDE and PD with or without agoraphobia, this revision should be none to very minor considering that the criteria changed very little or not at all. For other disorders, such as PTSD for which the criteria have undergone more changes, this revision may be more substantial. Therefore, comorbidity and the clusters that make up the various comorbidity structures will require reinvestigation as well, but we would expect little change for disorders for which the criteria have not changed considerably.

Validation of Comorbidity Factors

The 4 comorbidity factor scores, anxiety-depression, substance abuse, body image-eating, and depression-sleep problems were related to several variables: suicidal ideation, adequate social support, quality of life, self-confidence, and the Kessler-6 total score. As expected, the regressions of all 6 dependent variables on the 4 comorbidity factors were statistically significant. The greater the comorbidity factor scores the greater the Kessler-6 total score (the greater the psychological distress). A search of various psychological databases revealed very little about the association between psychological distress and current comorbidity other than 1 study which found a positive association between them [68]. Another study that was more focused on adults with seizures also found a positive relationship between psychiatric and physical comorbidities and psychological distress [76]. It should be noted that most of the variation in the Kessler-6 total score was accounted for by the anxiety-depression factor, which is expected because of the

direct relationship between Kessler-6 measures and measures of depression and anxiety.

As expected, the odds for having inadequate social support increased in likelihood for each additional point an individual scored on anxiety-depression, substance abuse, body image-eating, and depression-sleep problems factors, respectively. This is indeed expected because social relationships (involving social network composition, social support, social interaction frequency and quality, and the experience of isolation and loneliness) have long been linked with longevity and mental and physical health. Social support helps people to cope more adaptively with acute and chronic stress, thus potentially enhancing allostasis and autonomic nervous system (ANS) reactivity and related emotional regulation. Poor ANS reactivity and emotional dysregulation are generally considered core factors associated with many mental health conditions. Mazzella et al [71] concluded that mortality progressively increased with low social support and greater comorbidity, and that low social support progressively increased with increasing comorbidity in the elderly. Rockhill et al [72] found lack of social support mediated the association between symptoms and lower grades for adolescents with depression alone and comorbid symptoms. We did not find many studies that examined the direct relationship between social support and current comorbidity.

As expected, the odds for having low self-confidence to deal with one's mental health issues increased in likelihood for each additional point an individual scored on anxiety-depression, substance abuse, body image-eating, and depression-sleep problems factors, respectively. These results suggest that greater comorbidity factor scores reduce self-confidence in one's ability to deal with mental health issues. We are not aware of any study that directly examined the relationship between comorbidity of psychological disorders and patients' perceived confidence in dealing with their mental health issues.

As expected, the odds for having poor quality of life increased in likelihood for each additional point an individual scored on anxiety-depression, substance abuse, body image-eating, and depression-sleep problems factors, respectively. Several studies have shown a reduced health-related quality of life for patients with physical comorbidities [76-79]. Fewer studies have also found a negative relationship between comorbidity and quality of life in specific populations, such as obese or bipolar patients [69,70]. But we are not aware of any study that directly

examined psychological comorbidity and general quality of life. It should not be surprising to find a negative relationship between comorbidity factor scores and the general rating of one's quality of life. The same mechanisms that result in reduced quality of life for patients with physical comorbidities should also produce reduced quality of life for those with psychological comorbidity.

As expected, the odds for experiencing suicidal ideation increased in likelihood for each additional point an individual scored on anxiety-depression, substance abuse, body image-eating, and depression-sleep problems factors, respectively. These results are particularly important and useful because they suggest that comorbidity factor scores on all scales have the potential to be used for the purpose of identifying those with suicidal ideation, especially for those with high scores on the anxiety-depression factor. However, more work is needed to validate the potential use of these factor scores. These results are consistent with the findings of Albert et al [3] and Schoevers et al [4] who argued that comorbidity was consistently related to higher rate of suicide.

There are very few studies published on the relationships between current psychological comorbidity and measures such as general psychological distress, self-confidence in dealing with one's mental health, adequate social support, and quality of life. Future studies should further explore the relationships between these measures. The fact that we used only 1 simple question to measure these variables is a limitation of this study that places a constraint on the generalization of findings. A second limitation is that the e-PASS depends exclusively on automated self-report to assess and determine the diagnoses of participants. Andersson and Titov [80] raised some concerns about the use of an online assessment tool for the purpose of diagnosing individuals. Thirdly, there is only 1 study that examined the psychometric properties of the e-PASS (D Nguyen, unpublished PhD thesis, Swinburne University, 2013). More studies with larger samples and more established measures for comparisons based on the newly released *DSM-5* are needed before definitive conclusions can be made. A fourth limitation is that the 4 factors only managed to explain 44% of the total variance in the severity scores. This analysis needs to be repeated and a confirmatory factor analysis may be conducted on a new dataset collected since October 2012 to demonstrate the generalizability of this model.

Acknowledgments

The authors would like to thank the Australian Government Department of Health and Ageing who provided the funding for the development of the Anxiety Online service. e-PASS and the Anxiety Online e-therapy treatment programs are available to the international public at no cost (fully automated self-help). The therapist-assisted version is open only to Australian residents at a minimal cost. The funders of Anxiety Online (Australian Government Department of Health and Ageing) had no other involvement in this study or report.

Conflicts of Interest

None declared.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):593-602. [doi: [10.1001/archpsyc.62.6.593](https://doi.org/10.1001/archpsyc.62.6.593)] [Medline: [15939837](https://pubmed.ncbi.nlm.nih.gov/15939837/)]
2. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):617-627 [FREE Full text] [doi: [10.1001/archpsyc.62.6.617](https://doi.org/10.1001/archpsyc.62.6.617)] [Medline: [15939839](https://pubmed.ncbi.nlm.nih.gov/15939839/)]
3. Albert U, Rosso G, Maina G, Bogetto F. Impact of anxiety disorder comorbidity on quality of life in euthymic bipolar disorder patients: differences between bipolar I and II subtypes. *J Affect Disord* 2008 Jan;105(1-3):297-303. [doi: [10.1016/j.jad.2007.05.020](https://doi.org/10.1016/j.jad.2007.05.020)] [Medline: [17617468](https://pubmed.ncbi.nlm.nih.gov/17617468/)]
4. Schoevers RA, Deeg DJ, van Tilburg W, Beekman AT. Depression and generalized anxiety disorder: co-occurrence and longitudinal patterns in elderly patients. *Am J Geriatr Psychiatry* 2005 Jan;13(1):31-39. [doi: [10.1176/appi.ajgp.13.1.31](https://doi.org/10.1176/appi.ajgp.13.1.31)] [Medline: [15653938](https://pubmed.ncbi.nlm.nih.gov/15653938/)]
5. Slade T, Watson D. The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychol Med* 2006 Nov;36(11):1593-1600. [doi: [10.1017/S0033291706008452](https://doi.org/10.1017/S0033291706008452)] [Medline: [16882356](https://pubmed.ncbi.nlm.nih.gov/16882356/)]
6. Vollebergh WA, Iedema J, Bijl RV, de Graaf R, Smit F, Ormel J. The structure and stability of common mental disorders: the NEMESIS study. *Arch Gen Psychiatry* 2001 Jun;58(6):597-603. [Medline: [11386990](https://pubmed.ncbi.nlm.nih.gov/11386990/)]
7. Krueger RF. The structure of common mental disorders. *Arch Gen Psychiatry* 1999 Oct;56(10):921-926. [Medline: [10530634](https://pubmed.ncbi.nlm.nih.gov/10530634/)]
8. Wittchen HU, Höfler M, Merikangas K. Toward the identification of core psychopathological processes? *Arch Gen Psychiatry* 1999 Oct;56(10):929-931. [Medline: [10530635](https://pubmed.ncbi.nlm.nih.gov/10530635/)]
9. Beesdo-Baum K, Höfler M, Gloster AT, Klotsche J, Lieb R, Beauducel A, et al. The structure of common mental disorders: a replication study in a community sample of adolescents and young adults. *Int J Methods Psychiatr Res* 2009 Dec;18(4):204-220. [doi: [10.1002/mpr.293](https://doi.org/10.1002/mpr.293)] [Medline: [20024895](https://pubmed.ncbi.nlm.nih.gov/20024895/)]
10. Cox BJ, Clara IP, Enns MW. Posttraumatic stress disorder and the structure of common mental disorders. *Depress Anxiety* 2002;15(4):168-171. [doi: [10.1002/da.10052](https://doi.org/10.1002/da.10052)] [Medline: [12112721](https://pubmed.ncbi.nlm.nih.gov/12112721/)]
11. de Carvalho HW, Andreoli SB, Vaidyanathan U, Patrick CJ, Quintana MI, Jorge MR. The structure of common mental disorders in incarcerated offenders. *Compr Psychiatry* 2013 Feb;54(2):111-116. [doi: [10.1016/j.comppsy.2012.07.006](https://doi.org/10.1016/j.comppsy.2012.07.006)] [Medline: [22998843](https://pubmed.ncbi.nlm.nih.gov/22998843/)]
12. Oleski J, Cox BJ, Clara I, Hills A. Pathological gambling and the structure of common mental disorders. *J Nerv Ment Dis* 2011 Dec;199(12):956-960. [doi: [10.1097/NMD.0b013e3182392931](https://doi.org/10.1097/NMD.0b013e3182392931)] [Medline: [22134454](https://pubmed.ncbi.nlm.nih.gov/22134454/)]
13. McGlinchey JB, Zimmerman M. Examining a dimensional representation of depression and anxiety disorders' comorbidity in psychiatric outpatients with item response modeling. *J Abnorm Psychol* 2007 Aug;116(3):464-474. [doi: [10.1037/0021-843X.116.3.464](https://doi.org/10.1037/0021-843X.116.3.464)] [Medline: [17696702](https://pubmed.ncbi.nlm.nih.gov/17696702/)]
14. Naragon-Gainey K. Meta-analysis of the relations of anxiety sensitivity to the depressive and anxiety disorders. *Psychol Bull* 2010 Jan;136(1):128-150. [doi: [10.1037/a0018055](https://doi.org/10.1037/a0018055)] [Medline: [20063929](https://pubmed.ncbi.nlm.nih.gov/20063929/)]
15. Brady EU, Kendall PC. Comorbidity of anxiety and depression in children and adolescents. *Psychol Bull* 1992 Mar;111(2):244-255. [Medline: [1557475](https://pubmed.ncbi.nlm.nih.gov/1557475/)]
16. Gum AM, Cheavens JS. Psychiatric comorbidity and depression in older adults. *Curr Psychiatry Rep* 2008 Feb;10(1):23-29. [Medline: [18269891](https://pubmed.ncbi.nlm.nih.gov/18269891/)]
17. Fava M, Rankin MA, Wright EC, Alpert JE, Nierenberg AA, Pava J, et al. Anxiety disorders in major depression. *Compr Psychiatry* 2000;41(2):97-102. [Medline: [10741886](https://pubmed.ncbi.nlm.nih.gov/10741886/)]
18. Fava M, Alpert JE, Carmin CN, Wisniewski SR, Trivedi MH, Biggs MM, et al. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D. *Psychol Med* 2004 Oct;34(7):1299-1308. [Medline: [15697056](https://pubmed.ncbi.nlm.nih.gov/15697056/)]
19. Mackenzie CS, Reynolds K, Chou KL, Pagura J, Sareen J. Prevalence and correlates of generalized anxiety disorder in a national sample of older adults. *Am J Geriatr Psychiatry* 2011 Apr;19(4):305-315. [doi: [10.1097/JGP.0b013e318202bc62](https://doi.org/10.1097/JGP.0b013e318202bc62)] [Medline: [21427639](https://pubmed.ncbi.nlm.nih.gov/21427639/)]
20. Petersen T, Andreotti CF, Chelminski I, Young D, Zimmerman M. Do comorbid anxiety disorders impact treatment planning for outpatients with major depressive disorder? *Psychiatry Res* 2009 Aug 30;169(1):7-11. [doi: [10.1016/j.psychres.2008.10.036](https://doi.org/10.1016/j.psychres.2008.10.036)] [Medline: [19596155](https://pubmed.ncbi.nlm.nih.gov/19596155/)]
21. Atalay H. Comorbidity of insomnia detected by the Pittsburgh sleep quality index with anxiety, depression and personality disorders. *Isr J Psychiatry Relat Sci* 2011;48(1):54-59 [FREE Full text] [Medline: [21572244](https://pubmed.ncbi.nlm.nih.gov/21572244/)]
22. Taylor DJ, Lichstein KL, Durrence HH, Reidel BW, Bush AJ. Epidemiology of insomnia, depression, and anxiety. *Sleep* 2005 Nov;28(11):1457-1464. [Medline: [16335332](https://pubmed.ncbi.nlm.nih.gov/16335332/)]
23. Liu X, Buysse DJ, Gentzler AL, Kiss E, Mayer L, Kapornai K, et al. Insomnia and hypersomnia associated with depressive phenomenology and comorbidity in childhood depression. *Sleep* 2007 Jan;30(1):83-90. [Medline: [17310868](https://pubmed.ncbi.nlm.nih.gov/17310868/)]
24. Tsuno N, Besset A, Ritchie K. Sleep and depression. *J Clin Psychiatry* 2005 Oct;66(10):1254-1269. [Medline: [16259539](https://pubmed.ncbi.nlm.nih.gov/16259539/)]
25. Ivanenko A, Crabtree VM, Gozal D. Sleep and depression in children and adolescents. *Sleep Med Rev* 2005 Apr;9(2):115-129. [doi: [10.1016/j.smrv.2004.09.006](https://doi.org/10.1016/j.smrv.2004.09.006)] [Medline: [15737790](https://pubmed.ncbi.nlm.nih.gov/15737790/)]

26. Farabaugh A, Fava M, Mischoulon D, Sklarsky K, Petersen T, Alpert J. Relationships between major depressive disorder and comorbid anxiety and personality disorders. *Compr Psychiatry* 2005;46(4):266-271. [Medline: [16175757](#)]
27. Zimmerman M, McDermut W, Mattia JI. Frequency of anxiety disorders in psychiatric outpatients with major depressive disorder. *Am J Psychiatry* 2000 Aug;157(8):1337-1340. [Medline: [10910803](#)]
28. Diaconu G, Turecki G. Panic disorder and suicidality: is comorbidity with depression the key? *J Affect Disord* 2007 Dec;104(1-3):203-209. [doi: [10.1016/j.jad.2007.03.006](#)] [Medline: [17442399](#)]
29. Watson D. Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *J Abnorm Psychol* 2005 Nov;114(4):522-536. [doi: [10.1037/0021-843X.114.4.522](#)] [Medline: [16351375](#)]
30. Brown TA, Barlow DH. Comorbidity among anxiety disorders: implications for treatment and DSM-IV. *J Consult Clin Psychol* 1992 Dec;60(6):835-844. [Medline: [1460147](#)]
31. Andrews G, Anderson TM, Slade T, Sunderland M. Classification of anxiety and depressive disorders: problems and solutions. *Depress Anxiety* 2008;25(4):274-281. [doi: [10.1002/da.20489](#)] [Medline: [18415950](#)]
32. Clark LA, Watson D. Distress and fear disorders: an alternative empirically based taxonomy of the "mood" and "anxiety" disorders. *Br J Psychiatry* 2006 Dec;189:481-483 [FREE Full text] [doi: [10.1192/bjp.bp.106.03825](#)] [Medline: [17139030](#)]
33. Watson D, O'Hara MW, Stuart S. Hierarchical structures of affect and psychopathology and their implications for the classification of emotional disorders. *Depress Anxiety* 2008;25(4):282-288. [doi: [10.1002/da.20496](#)] [Medline: [18415951](#)]
34. Beesdo K, Pine DS, Lieb R, Wittchen HU. Incidence and risk patterns of anxiety and depressive disorders and categorization of generalized anxiety disorder. *Arch Gen Psychiatry* 2010 Jan;67(1):47-57. [doi: [10.1001/archgenpsychiatry.2009.177](#)] [Medline: [20048222](#)]
35. Marshall JR. Comorbidity and its effects on panic disorder. *Bull Menninger Clin* 1996;60(2 Suppl A):A39-A53. [Medline: [8857426](#)]
36. Mennin DS, Heimberg RG. The impact of comorbid mood and personality disorders in the cognitive-behavioral treatment of panic disorder. *Clin Psychol Rev* 2000 Apr;20(3):339-357. [Medline: [10779898](#)]
37. Welkowitz J, Welkowitz LA, Struening E, Hellman F, Guardino M. Panic and Comorbid Anxiety Symptoms in a National Anxiety Screening Sample: Implications for Clinical Interventions. *Psychotherapy: Theory, Research, Practice, Training* 2004;41(1):69-75. [doi: [10.1037/0033-3204.41.1.69](#)]
38. Casper RC. Depression and eating disorders. *Depress Anxiety* 1998;8 Suppl 1:96-104. [Medline: [9809221](#)]
39. Hudson JI, Pope HG, Yurgelun-Todd D, Jonas JM, Frankenburg FR. A controlled study of lifetime prevalence of affective and other psychiatric disorders in bulimic outpatients. *Am J Psychiatry* 1987 Oct;144(10):1283-1287. [Medline: [3499087](#)]
40. Iwasaki Y, Matsunaga H, Kiriike N, Tanaka H, Matsui T. Comorbidity of axis I disorders among eating-disordered subjects in Japan. *Compr Psychiatry* 2000;41(6):454-460. [doi: [10.1053/comp.2000.16561](#)] [Medline: [11086152](#)]
41. Jordan J, Joyce PR, Carter FA, Horn J, McIntosh VV, Luty SE, et al. Specific and nonspecific comorbidity in anorexia nervosa. *Int J Eat Disord* 2008 Jan;41(1):47-56. [doi: [10.1002/eat.20463](#)] [Medline: [17868127](#)]
42. Salbach-Andrae H, Lenz K, Simmendinger N, Klinkowski N, Lehmkuhl U, Pfeiffer E. Psychiatric comorbidities among female adolescents with anorexia nervosa. *Child Psychiatry Hum Dev* 2008 Sep;39(3):261-272. [doi: [10.1007/s10578-007-0086-1](#)] [Medline: [17987378](#)]
43. Blinder BJ, Cumella EJ, Sanathara VA. Psychiatric comorbidities of female inpatients with eating disorders. *Psychosom Med* 2006;68(3):454-462. [doi: [10.1097/01.psy.0000221254.77675.f5](#)] [Medline: [16738079](#)]
44. Godart NT, Flament MF, Perdereau F, Jeammet P. Comorbidity between eating disorders and anxiety disorders: a review. *Int J Eat Disord* 2002 Nov;32(3):253-270. [doi: [10.1002/eat.10096](#)] [Medline: [12210640](#)]
45. Pallister E, Waller G. Anxiety in the eating disorders: understanding the overlap. *Clin Psychol Rev* 2008 Mar;28(3):366-386. [doi: [10.1016/j.cpr.2007.07.001](#)] [Medline: [17707562](#)]
46. Swinbourne JM, Touyz SW. The co-morbidity of eating disorders and anxiety disorders: a review. *Eur Eat Disord Rev* 2007 Jul;15(4):253-274. [doi: [10.1002/erv.784](#)] [Medline: [17676696](#)]
47. Gadalla TM. Psychiatric comorbidity in eating disorders: a comparison of men and women. *Journal of Men's Health* 2008 Sep;5(3):209-217. [doi: [10.1016/j.jomh.2008.06.033](#)]
48. Godart NT, Flament MF, Lecrubier Y, Jeammet P. Anxiety disorders in anorexia nervosa and bulimia nervosa: co-morbidity and chronology of appearance. *Eur Psychiatry* 2000 Feb;15(1):38-45. [Medline: [10713801](#)]
49. Kaye WH, Bulik CM, Thornton L, Barbarich N, Masters K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *Am J Psychiatry* 2004 Dec;161(12):2215-2221. [doi: [10.1176/appi.ajp.161.12.2215](#)] [Medline: [15569892](#)]
50. Touchette E, Henegar A, Godart NT, Pryor L, Falissard B, Tremblay RE, et al. Subclinical eating disorders and their comorbidity with mood and anxiety disorders in adolescent girls. *Psychiatry Res* 2011 Jan 30;185(1-2):185-192. [doi: [10.1016/j.psychres.2010.04.005](#)] [Medline: [20546924](#)]
51. Godart NT, Flament MF, Curt F, Perdereau F, Lang F, Venisse JL, et al. Anxiety disorders in subjects seeking treatment for eating disorders: a DSM-IV controlled study. *Psychiatry Res* 2003 Mar 25;117(3):245-258. [Medline: [12686367](#)]
52. Hinrichsen H, Wright F, Waller G, Meyer C. Social anxiety and coping strategies in the eating disorders. *Eat Behav* 2003 Aug;4(2):117-126. [doi: [10.1016/S1471-0153\(03\)00016-3](#)] [Medline: [15000975](#)]
53. Keel PK, Klump KL, Miller KB, McGue M, Iacono WG. Shared transmission of eating disorders and anxiety disorders. *Int J Eat Disord* 2005 Sep;38(2):99-105. [doi: [10.1002/eat.20168](#)] [Medline: [16134107](#)]

54. Lilenfeld LR, Kaye WH, Greeno CG, Merikangas KR, Plotnicov K, Pollice C, et al. A controlled family study of anorexia nervosa and bulimia nervosa: psychiatric disorders in first-degree relatives and effects of proband comorbidity. *Arch Gen Psychiatry* 1998 Jul;55(7):603-610. [Medline: [9672050](#)]
55. Allen A, Hollander E. Similarities and differences between body dysmorphic disorder and other disorders. *Psychiatric Annals* 2004;34(12):927-933.
56. Cororve MB, Gleaves DH. Body dysmorphic disorder: a review of conceptualizations, assessment, and treatment strategies. *Clin Psychol Rev* 2001 Aug;21(6):949-970. [Medline: [11497214](#)]
57. Rauch SL, Phillips KA, Segal E, Makris N, Shin LM, Whalen PJ, et al. A preliminary morphometric magnetic resonance imaging study of regional brain volumes in body dysmorphic disorder. *Psychiatry Res* 2003 Jan 20;122(1):13-19. [Medline: [12589879](#)]
58. Phillips KA. *The Broken Mirror: Understanding and Treating Body Dysmorphic Disorder*. Oxford: Oxford University Press; 2005.
59. Conway KP, Compton W, Stinson FS, Grant BF. Lifetime comorbidity of DSM-IV mood and anxiety disorders and specific drug use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Clin Psychiatry* 2006 Feb;67(2):247-257. [Medline: [16566620](#)]
60. Rush B, Koegl CJ. Prevalence and profile of people with co-occurring mental and substance use disorders within a comprehensive mental health system. *Can J Psychiatry* 2008 Dec;53(12):810-821. [Medline: [19087479](#)]
61. Teesson M, Slade T, Mills K. Comorbidity in Australia: findings of the 2007 National Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry* 2009 Jul;43(7):606-614. [doi: [10.1080/00048670902970908](#)] [Medline: [19530017](#)]
62. McGovern MP, Xie H, Segal SR, Siembab L, Drake RE. Addiction treatment services and co-occurring disorders: Prevalence estimates, treatment practices, and barriers. *J Subst Abuse Treat* 2006 Oct;31(3):267-275. [doi: [10.1016/j.jsat.2006.05.003](#)] [Medline: [16996389](#)]
63. Katz LY, Cox BJ, Clara IP, Oleski J, Sacevich T. Substance abuse versus dependence and the structure of common mental disorders. *Compr Psychiatry* 2011;52(6):638-643. [doi: [10.1016/j.comppsy.2010.12.001](#)] [Medline: [21295775](#)]
64. Slade T, Grove R, Teesson M. A taxometric study of alcohol abuse and dependence in a general population sample: evidence of dimensional latent structure and implications for DSM-V. *Addiction* 2009 May;104(5):742-751. [doi: [10.1111/j.1360-0443.2009.02498.x](#)] [Medline: [19215598](#)]
65. Klein B, Meyer D, Austin DW, Kyrios M. Anxiety online: a virtual clinic: preliminary outcomes following completion of five fully automated treatment programs for anxiety disorders and symptoms. *J Med Internet Res* 2011;13(4):e89 [FREE Full text] [doi: [10.2196/jmir.1918](#)] [Medline: [22057287](#)]
66. AL-Asadi AM, Klein B, Meyer D. Posttreatment Attrition and Its Predictors, Attrition Bias, and Treatment Efficacy of the Anxiety Online Programs. *J Med Internet Res* 2014;16(10):e232 [FREE Full text] [doi: [10.2196/jmir.3513](#)]
67. National eTherapy Centre. Mental Health Online. URL: <https://www.anxietyonline.org.au/default.aspx> [accessed 2014-10-18] [WebCite Cache ID 6TQZHUnGL]
68. Andrews G, Slade T, Issakidis C. Deconstructing current comorbidity: data from the Australian National Survey of Mental Health and Well-Being. *Br J Psychiatry* 2002 Oct;181:306-314 [FREE Full text] [Medline: [12356657](#)]
69. Carpiniello B, Pinna F, Pillai G, Nonnoi V, Pisano E, Corrias S, et al. Psychiatric comorbidity and quality of life in obese patients. Results from a case-control study. *Int J Psychiatry Med* 2009;39(1):63-78. [Medline: [19650530](#)]
70. Kauer-Sant'Anna M, Frey BN, Andreatza AC, Ceresér KM, Gazalle FK, Tramontina J, et al. Anxiety comorbidity and quality of life in bipolar disorder patients. *Can J Psychiatry* 2007 Mar;52(3):175-181. [Medline: [17479526](#)]
71. Mazzella F, Cacciatore F, Galizia G, Della-Morte D, Rossetti M, Abbruzzese R, et al. Social support and long-term mortality in the elderly: role of comorbidity. *Arch Gerontol Geriatr* 2010;51(3):323-328. [doi: [10.1016/j.archger.2010.01.011](#)] [Medline: [20153534](#)]
72. Rockhill C, Vander Stoep A, McCauley E, Katon W. Social competence and social support as mediators between comorbid depressive and conduct problems and functional outcomes in middle school children. *J Adolesc* 2009 Jun;32(3):535-553 [FREE Full text] [doi: [10.1016/j.adolescence.2008.06.011](#)] [Medline: [18694594](#)]
73. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 2003 Feb;60(2):184-189. [Medline: [12578436](#)]
74. Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika* 1965 Jun;30:179-185. [Medline: [14306381](#)]
75. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behav Res Methods Instrum Comput* 2000 Aug;32(3):396-402 [FREE Full text] [Medline: [11029811](#)]
76. Strine TW, Kobau R, Chapman DP, Thurman DJ, Price P, Balluz LS. Psychological distress, comorbidities, and health behaviors among U.S. adults with seizures: results from the 2002 National Health Interview Survey. *Epilepsia* 2005 Jul;46(7):1133-1139. [doi: [10.1111/j.1528-1167.2005.01605.x](#)] [Medline: [16026567](#)]
77. Zyguntowicz M, Owczarek A, Elibol A, Chudek J. Comorbidities and the quality of life in hypertensive patients. *Pol Arch Med Wewn* 2012;122(7-8):333-340 [FREE Full text] [Medline: [22814517](#)]

78. Haagsma JA, van Beeck EF, Polinder S, Toet H, Panneman M, Bonsel GJ. The effect of comorbidity on health-related quality of life for injury patients in the first year following injury: comparison of three comorbidity adjustment approaches. *Popul Health Metr* 2011;9:10 [FREE Full text] [doi: [10.1186/1478-7954-9-10](https://doi.org/10.1186/1478-7954-9-10)] [Medline: [21513572](https://pubmed.ncbi.nlm.nih.gov/21513572/)]
79. Bayliss EA, Ellis JL, Steiner JF. Subjective assessments of comorbidity correlate with quality of life health outcomes: initial validation of a comorbidity assessment instrument. *Health Qual Life Outcomes* 2005;3:51 [FREE Full text] [doi: [10.1186/1477-7525-3-51](https://doi.org/10.1186/1477-7525-3-51)] [Medline: [16137329](https://pubmed.ncbi.nlm.nih.gov/16137329/)]
80. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry* 2014 Feb;13(1):4-11 [FREE Full text] [doi: [10.1002/wps.20083](https://doi.org/10.1002/wps.20083)] [Medline: [24497236](https://pubmed.ncbi.nlm.nih.gov/24497236/)]

Abbreviations

ADIS: Anxiety Disorders Interview Schedule

ANS: autonomic nervous system

BDD: body dysmorphic disorder

DSM-5: Diagnostic and Statistical Manual of Mental Disorders (5th Edition)

DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders (4th Edition, Text Revision)

e-PASS: electronic psychological assessment screening system

GAD: generalized anxiety disorder

KMO: Kaiser-Meyer-Olkin

MDD: major depressive disorder

MDE: major depressive episode

NCS: National Comorbidity Survey

OCD: obsessive-compulsive disorder

PCA: principal component analysis

PD: panic disorder

PTSD: posttraumatic stress disorder

SAD: social anxiety disorder

Edited by G Eysenbach; submitted 09.06.14; peer-reviewed by P Batterham; comments to author 30.06.14; revised version received 16.07.14; accepted 13.08.14; published 28.10.14.

Please cite as:

AL-Asadi AM, Klein B, Meyer D

Comorbidity Structure of Psychological Disorders in the Online e-PASS Data as Predictors of Psychosocial Adjustment Measures: Psychological Distress, Adequate Social Support, Self-Confidence, Quality of Life, and Suicidal Ideation

J Med Internet Res 2014;16(10):e248

URL: <http://www.jmir.org/2014/10/e248/>

doi: [10.2196/jmir.3591](https://doi.org/10.2196/jmir.3591)

PMID: [25351885](https://pubmed.ncbi.nlm.nih.gov/25351885/)

©Ali M AL-Asadi, Britt Klein, Denny Meyer. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

What Are Health-Related Users Tweeting? A Qualitative Content Analysis of Health-Related Users and Their Messages on Twitter

Joy L Lee¹, MS; Matthew DeCamp², MD, PhD; Mark Dredze³, PhD; Margaret S Chisolm⁴, MD; Zackary D Berger², MD, PhD

¹Johns Hopkins Bloomberg School of Public Health, Department of Health Policy & Management, Baltimore, MD, United States

²Johns Hopkins Berman Institute of Bioethics & Division of General Internal Medicine, Baltimore, MD, United States

³Johns Hopkins University, Human Language Technology Center of Excellence, Baltimore, MD, United States

⁴Johns Hopkins University School of Medicine, Department of Psychiatry and Behavioral Sciences, Baltimore, MD, United States

Corresponding Author:

Joy L Lee, MS

Johns Hopkins Bloomberg School of Public Health

Department of Health Policy & Management

624 North Broadway

Baltimore, MD, 21295

United States

Phone: 1 978 319 3618

Fax: 1 410 955 0470

Email: jllee@jhu.edu

Abstract

Background: Twitter is home to many health professionals who send messages about a variety of health-related topics. Amid concerns about physicians posting inappropriate content online, more in-depth knowledge about these messages is needed to understand health professionals' behavior on Twitter.

Objective: Our goal was to characterize the content of Twitter messages, specifically focusing on health professionals and their tweets relating to health.

Methods: We performed an in-depth content analysis of 700 tweets. Qualitative content analysis was conducted on tweets by health users on Twitter. The primary objective was to describe the general type of content (ie, health-related versus non-health related) on Twitter authored by health professionals and further to describe health-related tweets on the basis of the type of statement made. Specific attention was given to whether a tweet was personal (as opposed to professional) or made a claim that users would expect to be supported by some level of medical evidence (ie, a "testable" claim). A secondary objective was to compare content types among different users, including patients, physicians, nurses, health care organizations, and others.

Results: Health-related users are posting a wide range of content on Twitter. Among health-related tweets, 53.2% (184/346) contained a testable claim. Of health-related tweets by providers, 17.6% (61/346) were personal in nature; 61% (59/96) made testable statements. While organizations and businesses use Twitter to promote their services and products, patient advocates are using this tool to share their personal experiences with health.

Conclusions: Twitter users in health-related fields tweet about both testable claims and personal experiences. Future work should assess the relationship between testable tweets and the actual level of evidence supporting them, including how Twitter users—especially patients—interpret the content of tweets posted by health providers.

(*J Med Internet Res* 2014;16(10):e237) doi:[10.2196/jmir.3765](https://doi.org/10.2196/jmir.3765)

KEYWORDS

communication; consumer health informatics; health information technology; social media

Introduction

Close to 90% of US adults use the Internet [1]; 72% of those visit social media websites [2]. In fact, 12% of Internet users

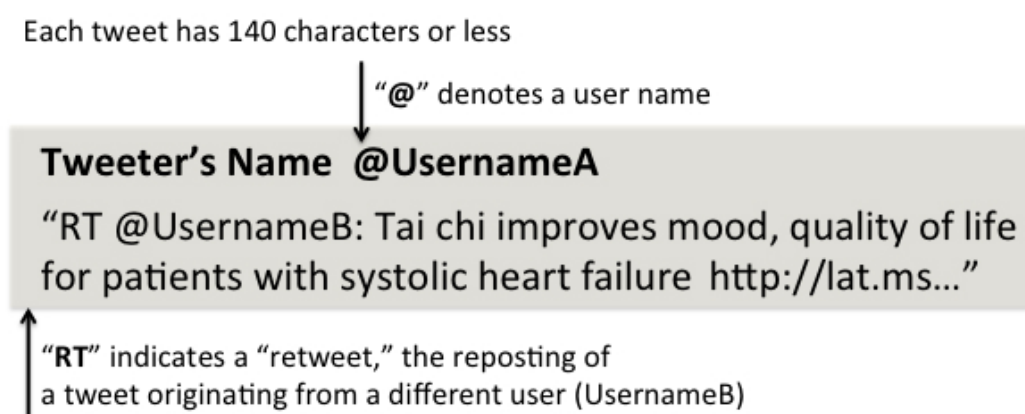
access social media to research health issues [3]. In particular, the micro-blogging service Twitter, which enables users to send short messages no more than 140 characters in length (see [Figure 1](#) for the breakdown of a tweet), has quickly grown in popularity,

with an estimated 255 million users monthly and 22% of active accounts based in the United States [4].

Health providers, patients, and patient advocates comment on a diverse array of topics on Twitter [5-7]. Previous reports have examined discrete health topics that users are discussing on Twitter [8-10], as well as concerns related to medical professionalism [11,12]. Health researchers have also used Twitter as a tool to track disease outbreak [13,14] and recruit study subjects [15]. However, little in-depth knowledge is available about health-related content by health professionals on Twitter, in particular whether such content represents personal opinion, a claim that those who view the tweet would expect to be supported by rigorous evidence, or something else. In addition, despite recent controversy over whether physicians should separate their “personal” and “professional” identities online [16-18], little is known regarding health professionals’ actual behavior.

Filling these knowledge gaps requires investigating message content at a depth only possible through qualitative analysis of health professionals’ tweets. This paper reports the results of a content analysis of such tweets. The primary objective of the study was to describe the general types of content on Twitter authored by health professionals, focusing on health-related content. Key content types of interest included content that would be perceived as “personal” and content that users would expect to be supported by medical evidence (ie, “testable” claims about health). Secondary objectives included comparisons of content areas by user type (eg, providers, health advocates). Based on prior experience and existing professional guidelines, it was hypothesized that self-identified health professionals would not post personal content on Twitter and that testable claims would be rare.

Figure 1. Anatomy of a tweet.



Methods

Data Collection

Data collection occurred in three basic steps. First, health users of Twitter were identified. Lists of Twitter users created by Organized Wisdom [19], a company in the health information technology sector, were used to gather data. Organized Wisdom collates health-related Twitter users into lists that enable others to more efficiently find information online; for example, the “doctors verified” list consists of physicians who openly disclose their credentials online. Of these lists, 37 are Twitter accounts for doctors, nurses, hospitals, and other health organizations that cover seven user groups of interest: health advocates, businesses, health care organizations, government organizations, students, professionals, and other (eg, health care-related publications).

Second, from these health users, an equal number (n=200) of the most recent publicly available tweets for each user in each group, including retweets, was downloaded during the period July 31-August 3, 2012. In cases where users had fewer than 200 tweets, all available data were obtained.

Third, from this large collection of tweets from the seven user groups, a smaller collection was sampled. For each group, a user from the group was randomly selected, and a random tweet (without replacement) was selected for that user. This strategy was used to create a preliminary sample of tweets (15 random tweets from each user group, or 105 total) to develop a coding scheme and a final sample for in-depth qualitative analysis (100 random tweets from each user group, or 700 total).

Content Analysis

The research team first hypothesized content areas of interest from their own experience and the published literature to generate a draft coding scheme. Then, using the preliminary sample of 105 tweets, 4 coauthors (JL, ZB, MD, MC) conducted content analyses of each tweet using this draft scheme. The coding scheme was then revised for the final analysis. The categories of interest are presented in Table 1. Subsequent analysis was then conducted on an additional 700 tweets as described above. Areas of disagreement between reviewer pairs were resolved by consensus in a case-by-case fashion. Reviewers were blind to the identity of the Twitter users during content analysis.

Table 1. Categories of health-related tweets identified via content analysis (n=346; content frequency may be greater than 100% because content areas were not mutually exclusive).

Tweet type	Definition	Example	Frequency among health tweets, %
Testable	An evidence-based claim whose veracity could be assessed	RT @username: Did you know? You can become more sensitive to alcohol as you get older: http://[link] #aging	53
News	Report of information or events that happened recently, often refers to traditional news outlet	“@username: State charges man with #stealing \$400,000 from church, #Alzheimer’s patient, choral group http://[link] #embezzle	41
Commercial	Refers to or advertises a service or product for sale	RT @username: Learn to capture/edit/produce quality videos at ASGE’s Video Editing Course for Physicians, June 8-10, Oak Brook, IL. ...	27
Wellness	Refers to food, diet, or exercise	RT @username: Ready for playoff football? Enjoy the wild card games & stick to a healthy diet with these recipes: http://[link]	14
Personal	Refers to personal experiences relayed by users	I visited @username today for the first time. Cool space but where are the women? Let’s get the women who made the @username pledge going!	18

After unmasking the data, users and user groups were reviewed to verify the accuracy of categories provided by Organized Wisdom. None of the tweets reviewed were found to be authored by the research team. In several instances, the provided categories were inaccurate; for example, a self-identified hotel was miscategorized as a government agency. Therefore, two members of the research team (JL and ZB) re-categorized all

users into seven new user groups based on consensus review of publicly available profiles (Table 2): users without available profiles, non-health users, health advocates, health businesses, non-provider health professionals, health providers, and health care organizations. This resulted in a different number of tweets for each of the seven groups.

Table 2. Types of users.

User type definition	User examples	Users (n=255), % (n)	Health tweets (n=346), n	Tweet examples
No profile: Accounts deleted or suspended after data collection so profiles could not be assessed	Not available	7 (19)	19	Vision Therapy Improves Vision Related Learning Problems: LIVONIA, MICHIGAN -- When children are having trouble... http://[link]
Non-health related users: Users unaffiliated with a health profession, organization, or mission	Justin Bieber fan account; a hotel in Gloucestershire, England; expert in “online promotion”	25 (64)	58	RT @username: These Baked Honey-Mustard Chicken Bites r big on flavor - and small on POINTS® Value! http://[link]
Health advocates: Users educating and advocating health issues on behalf of patients	Patient advocate with diabetes; educator and promoter on disabilities issues	6 (16)	26	My Life as Mandy... with Epilepsy: Educational Information http://[link]
Non-provider health professionals: Users working in the health field who do not provide direct patient care	Health economist; registered dietitian; medical student	11 (28)	43	Hey girl, I so hear ya! I’m on the same boat! RT @username: ABSOLUTELY NEEDS TO LOSE WEIGHT!! I CANT KEEP PROCRASTINATING! Get eht 2qethr.
Health providers: Direct-care providers, such as physicians and nurses	A registered ER nurse “and mother of 3”; a clinical cardiologist; “physician, keynote speaker, and media health expert”	22 (56)	59	Human factors: articulate video on patient safety inspired from personal loss http://[link]
Health businesses: Commercial businesses selling specific service or product	Herbal supplements company; health care marketing agency	22 (55)	99	@username Good luck with all that reading. Have you seen our Anatomy & Physiology Online? Makes learning easier http://[link]
Organization: Hospitals, medical societies, clinics, and journals	The Cancer Letter; Albert Einstein College of Medicine; Kids in Need of Dental Services	7 (17)	42	RT @username: Public #malaria drug research #data now more discoverable http://[link] #open #science20

Statistical Analysis

Descriptive statistics were used to tabulate types of tweets for each group of users. Statistically significant differences between the proportion of tweets among user groups were assessed using the chi-square test. Stata 13 was used for the analysis.

This study was reviewed and declared exempt from further review by the Johns Hopkins Medicine Institutional Review Board.

Results

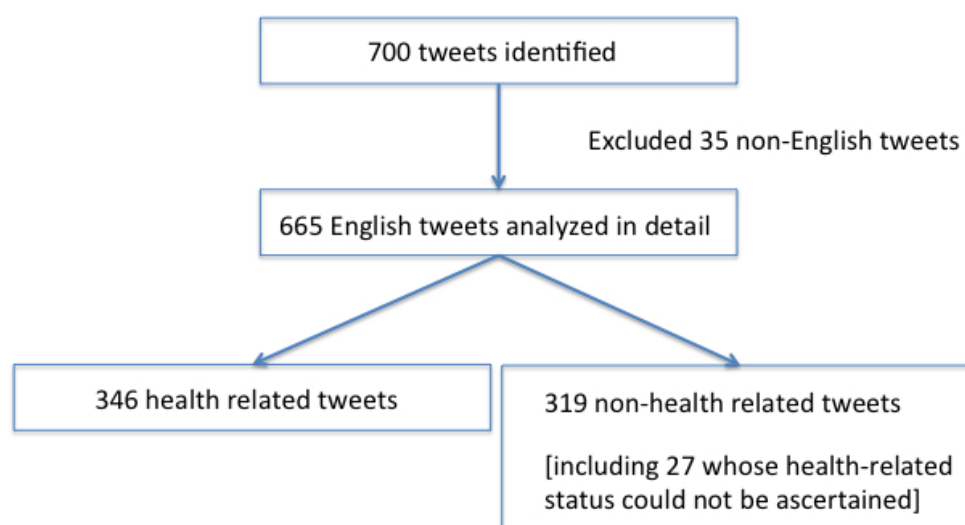
Tweets Overall

Of all tweets, 5.0% (35/700) non-English tweets were excluded. The remaining 95.0% (665/700) English-language tweets were

analyzed in depth (Figure 2). A total of 255 unique users contributed to the 665 tweets. Of the tweets in the analytic sample, 52.0% (346/665) were categorized as health-related. For example, a tweet such as “RT @Username: Tai chi improves mood, quality of life for patients with systolic heart failure” was categorized as health-related. Each health-related tweet was thereafter sub-coded into at least one of six non-mutually exclusive categories, as defined by Table 1.

Nearly one-third (31.1%, 207/665) of all tweets in the sample were personal in nature. Among personal tweets, 70.5% (146/207) were non-health related, whereas 29.5% (61/207) were health-related. This finding was statistically significant ($P<.001$).

Figure 2. Inclusion and exclusion of tweets.



Content Analysis of Health-Related Tweets

The research team was most interested in health-related tweets. Just over half (53.2%, 184/346) of these tweets contained “testable” claims, defined as claims for which someone viewing the tweet might expect to be supported by medical evidence. An example of a testable tweet is shown in Table 1 (first row).

Within health-related tweets, news was the second most frequent content category; 41.0% (142/346) of these tweets contained news. Over a quarter (26.9%, 93/346) contained reference to a commercial product or service, 17.6% (61/346) relayed a personal viewpoint or experience, and 17.1% (59/346) discussed wellness.

Categories of Health-Related Tweets by User Type

Overview

The analysis also compared the types of health-related tweets made by each of the seven user groups. Many of the users generating non-English tweets (9/14) contributed English tweets as well. The characterization of the health-related tweets by each user group is as follows.

No Profile

For 7.5% (19/255) of users, contributing 5.1% (34/665) of tweets, the user profile could not be accessed. Of these tweets, 56% (19/34) had health-related content. Among health-related tweets, 68% (13/19) made testable claims. Because 32% (6/19) of tweets by this group were personal, some of these users may have been individuals.

Non-Health-Related User

During the analysis, it was noted by the research team that—although users in the health field were originally intentionally sampled using Organized Wisdom—about a quarter of the users in the analysis (62/255) were not affiliated with a health profession, organization, or mission. This was the largest user group in the sample and potentially represents a comparison group of non-health users. These users contributed 30.2% (201/665) of all tweets in the analysis and 16.8% (58/346) of health-related tweets.

Non-health users tweeted more about wellness than health-related users: 33% vs health user average of 14%, $P<.001$; health users defined as all groups other than non-health-related users, including no profile, health providers, health-related businesses, non-provider health professionals, patient advocates, and organizations. More so than any other group, the majority

of tweets by this group also included a link (86%, 50/58), and significantly fewer of those links were accessible by the research team (43%, vs health user average of 80%, $P < .001$). Of all tweets by these users, 55% (32/58) had a testable claim and 47% (27/58) were news.

Health Providers

Direct-care providers, such as physicians and nurses, encompassed the next largest group of users, accounting for 23.1% (59/255) of the analytic sample. These users contributed 14.4% (96/665) of tweets analyzed; 61.5% (59/96) of tweets by these users were health-related in nature. They shared testable claims and news (61%, 36/59 and 56%, 33/59 respectively) more than other groups of individual users, although 14% (8/59) also mentioned a commercial product or service and 17% (10/59) were personal (eg, "That awkward moment when there is no awkward moment to tweet about").

Health-Related Businesses

About one-fifth of the users (21.6%, 55/255) in the analysis were health-related businesses whose health-related tweets constituted 28.6% (99/346) of the analytic sample. Many tweets by health businesses (44%, 44/99) mentioned a product or service, the highest percentage among the categories of users. Aside from these commercially oriented tweets, businesses often shared information seemingly relevant to their market (eg, "When You Leave Someone with a Mental Illness" from an addiction center or "#Pharma cos. push for social-media rules #hscm #hcmktg #web20" from a marketing agency). These users also had significantly more tweets that were personal in content (8% vs health user average of 19%, $P = .003$) than other health users.

Non-Provider Health Professionals

Non-provider health professionals are those in health care professions who do not provide direct patient care. Non-provider professionals made up 11.0% (28/255) of users in the sample, including two medical students; 12.4% (43/346) of total health tweets came from this group. The two medical students contributed nearly half of the tweets in this group (47%, 20/43). Only 49% (43/87) of tweets by this group of users were health-related, which was the lowest percentage among health users. This group of users made the fewest testable claims and referenced the news the least in their health tweets as compared to all other user groups. Instead, nearly half of the content (44%, 19/43) was about the users' personal experiences. Among the two medical students, 70% (14/20) of health tweets were personal in nature. At the same time, this group of users also tweeted less about commercial products or services compared to any other group.

Patient Advocates

Patient advocates made up 6.3% (16/255) of the users in our analysis, and together they contributed 6% (45/665) of tweets. More than half (58%, 26/45) of tweets by these users were health-related; 38% (10/26) of the users' health-related tweets were personal in nature, the second-highest number of personal tweets among the groups. Only 8% (2/26) of health tweets by this group touched upon wellness, which was below the average of 14%. Many of the messages retweeted content from other users, and 85% (22/26) included a link.

Organizations

Health care organizations represented in this sample were hospitals, medical societies, health care clinics, and journals. These users represented 5% of our sample and contributed 7.7% (51/665) of tweets. More than half (82%, 42/51) of their tweets were health-related. Moreover, these health tweets were most likely to share "testable" claims than other users. The majority of organizations' tweets contained testable claims (64%, 27/42) and news (52%, 22/42), and few were personal in nature (7%, 3/42). Yet organizations also had a high number of commercial tweets; a third of their health tweets (33%, 14/42) advertised a commercial product, or most often, their own services (eg, CPR courses, lab tests, and staff education curriculum).

Like tweets by patient advocates, tweets by health care organizations often passed on information through links. Most of their links (91%, 32/35) were accessible, compared to an overall average of 72% ($P = .04$).

Group Comparisons

Table 3 summarizes the distribution of tweet content areas among user groups, where content by health providers serves as the baseline for comparison. Overall, there were no significant differences in the types of content between health providers and non-health users and those without available profiles.

Health advocates tweeted about personal health matters significantly more than providers. Provider tweets were most different from those of health care businesses. Businesses were statistically more likely to tweet about news and less likely to tweet about wellness, and non-provider health professionals were statistically more likely to tweet testable claims and news stories and less about personal content. The tweets of health care organizations and providers differed as well. Organizations' tweets were statistically more likely to include commercial content and wellness but were less likely to include personal content.

Table 3. Percentage of health-related tweets in selected content areas by user group.

	Testable	News	Commercial	Wellness	Personal
Providers	61	56	14	25	17
No profile	68	32	16	10	32
Not health	55	47	22	33	9
Business	47	34 ^a	44 ^a	12 ^a	8
Non-Providers	37 ^a	26 ^a	16	12	44 ^a
Advocate	50	35	15	7	38 ^a
Organization	64	52	33 ^a	10 ^a	7 ^a

^aStatistically significant difference compared to the provider group at $P < .05$ using individual chi-square comparisons.

Discussion

Principal Results

This analysis of self-identified health users and their health-related messages on Twitter identified several important findings. First, over half of the tweets in the sample included tweets with testable claims, defined for this study as content that a Twitter user might expect to be supported by medical evidence. The majority of health tweets were testable and comprised various topics. In a preliminary analysis, we found that about 40% of such testable tweets shared medical advice while 23% disseminated scientific news or research books and articles. Although the research team did not analyze this in detail, the sources behind these medical claims were more varied than scientific research (ie, “evidence-based” medicine claims), ranging from common sense (eg, advice to drink more water on a hot day) to WebMD to Dr. Oz to personal blogs. Whether the claims were actually scientifically valid was not examined and requires further study.

Second, in addition to traditional providers, a wide range of non-health providers are sending health-related tweets. Moreover, the content of tweets by health-related users on Twitter varied with user type. Health care organizations, for example, shared more health news and testable claims than individuals but also endorsed commercial services or products more than individuals. In the future, it could be important to examine whether organizations’ tweets are based more often in rigorous evidence as compared to those of individual health professionals. These findings confirm and expand upon the work of Sugawara et al [20] and Chretien et al [21], who both focused on a subset of health users. In examining the tweets of cancer patients in Japan, Sugawara found that patients tweeted about both personal (eg, greetings, chats) and medical matters (eg, information about treatments). Similarly, Chretien found that health providers (specifically, physicians) shared personal, commercial, as well as medical information over Twitter.

We found that 17% of users in this sample were non-provider individuals in the health field, together contributing 20% of health-related tweets. Their participation illustrates the wide range of non-providers engaging in information sharing and health messaging on Twitter. While this community of users represents a significant resource of health conversation and

information on Twitter, these non-providers are not bound by professional guidelines on Twitter use. It is unclear how Twitter users might use or understand the tweets of this group compared, for example, to tweets by physicians and nurses. Given that trust is a significant feature of social networks (with young users potentially trusting their social network as much as physicians), this requires further examination [3,22].

Third, these findings demonstrate potential gaps between professional guidelines and online behavior [23,24]. Many medical societies, institutions, and medical schools disseminate such guidelines for clinicians and students. Often, these guidelines encourage the use of social media as a tool for endorsing medical information, discourage using the medium for medical consults, and counsel against posting personal content. Despite this, 17% of health tweets from providers and 74% from medical students included personal content. This lack of adherence to professional guidelines identifies an area where additional education or potentially consensus building within the profession is needed. Similarly, the sizable minority (14%) of providers’ tweets that included commercial products or services confirms prior findings and concerns about potential conflicts of interest on Twitter (a medium where disclosure is difficult, if not impossible) [21,25].

As the use of Twitter has increased, so has the scholarship on the use of Twitter for health-related purposes. Some researchers have used Twitter as a surveillance [7,13,26] and prediction tool for disease outbreak [27] while others have characterized the relationship between segments of health-related users on Twitter, such as medical societies [28], internal medicine physicians [29], and cancer patients [20]. Though few papers examined health-related users as a whole, or the content of their tweets in detail, ours is not the first content analysis of Twitter. For example, Chew and Eysenbach analyzed individual tweets relating to the 2009 H1N1 pandemic, finding that information from a variety of credible sources, in addition to opinions and experiences, were disseminated on Twitter [26]. Our study, a qualitative analysis of tweets from health-related users, addressed different questions and thus made use of a different approach.

This study adds to the growing literature on using Twitter for health purposes by examining the content of tweets by health-related users. In doing so, this study had several important strengths. It represents an in-depth exploration of the types of

health-related users and content on Twitter. The analysis is the first to examine in detail both health users and their health-related content on Twitter, orienting researchers and clinicians to the Twitter health landscape. While previous studies focused on Twitter use among certain populations or users [8,20,21,30,31] such as cancer patients in Japan, or certain topics such as childhood obesity [9], this study explored several types of users, achieving breadth and qualitative depth simultaneously.

Limitations

The study was not without limitations. Data collection occurred in the summer of 2012; this sample thus reflects a Twitter landscape that may have shifted since then. For example, adherence to professional guidelines could be expected to increase over time. This study's findings set a baseline for future analyses. The selection of the analytic sample was limited by the absence of a simple mechanism to sample health-related users and tweets on Twitter. For this reason, it was not feasible to select a random sample of all health users and their tweets, reducing the generalizability of the study findings. Although Organized Wisdom, by organizing content into lists, seemed to offer a reasonable way to address this problem, many user category designations were inaccurate. Favoring accuracy of user type over preservation of the original categories, the research team undertook its own characterizations. Because this reduced the number of tweets in certain user groups, this reduced the ability to determine statistical significance between groups. If so, the results likely underestimate potential differences between groups. Importantly, this only reinforces the need for in-depth analysis while revealing a concern that collated "health-related user" content might not be from health-related users.

The analysis did not purposely evaluate "high impact" users (eg, Dr. Oz, Atul Gawande, Sanjay Gupta). We also did not independently verify user profiles, and users representing themselves as providers were not checked against other databases. Furthermore, while the analysis was double coded and disagreements were resolved by consensus, the subjective nature of coding limits the validity of these findings.

Future Considerations

This study presents several opportunities for further research. While a preliminary assessment evaluated the evidence base of each testable tweet, future analysis could examine this subset of tweets (health-related tweets that contain testable claims), for example, by comparing them to accepted strength of evidence standards [32]. Cross-referencing the identities of the users would be another next step in refining the characterization of users, while including Twitter users who are verifiable, self-identified patients or providers would further enrich the analysis of health messages.

Conclusions

In conclusion, health-related users and the content they share on Twitter are diverse. While providers' tweets often include testable claims, they also make use of personal statements. Non-providers also contribute to ongoing health-related messaging on Twitter. Users in health-related fields who are actively engaging in health-related conversations on Twitter, as well as those who are merely reading these users' tweets, should be oriented to the diversity of health-related Twitter content, and, if possible, to the validity of associated tweets.

Conflicts of Interest

Dr DeCamp reports having received travel reimbursement from the American College of Physicians related to speaking on social media and medical professionalism. Dr Berger reports receiving compensation from the Johns Hopkins University School of Medicine for generating tweets as part of a medical education website.

References

1. Pew Research Center. Pew Research Center's Internet and American Life Project. 2014 Jan. Internet User Demographics URL: <http://www.pewinternet.org/data-trend/internet-use/latest-stats/> [accessed 2014-10-01] [WebCite Cache ID 6T05cSF65]
2. Brenner J, Smith A. Pew Research Center Internet & American Life Project Tracking Survey. 2013 Apr. 72% of online adults are social networking site users URL: <http://www.pewinternet.org/2013/08/05/72-of-online-adults-are-social-networking-site-users/> [accessed 2014-10-01] [WebCite Cache ID 6T05jP2sH]
3. Fox S, Jones S. Pew Internet & American Life Project Survey. 2009. The social life of health information Internet URL: http://www.pewinternet.org/files/old-media/Files/Reports/2009/PIP_Health_2009.pdf [accessed 2014-10-01] [WebCite Cache ID 6T08AHdFO]
4. Twitter. About page URL: <https://about.twitter.com/company> [accessed 2014-10-01] [WebCite Cache ID 6T08Egbe]
5. Paul MJ, Dredze M. You are what you Tweet: Analyzing Twitter for public health. 2011 Presented at: International Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona, Spain.
6. Mehta N, Flickinger T. The Times They Are A-Changin': Academia, Social Media and the JGIM Twitter Journal Club. *J Gen Intern Med* 2014 Oct;29(10):1317-1318. [doi: [10.1007/s11606-014-2976-9](https://doi.org/10.1007/s11606-014-2976-9)] [Medline: [25092011](https://pubmed.ncbi.nlm.nih.gov/25092011/)]
7. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. *Drug Saf* 2014 May;37(5):343-350 [FREE Full text] [doi: [10.1007/s40264-014-0155-x](https://doi.org/10.1007/s40264-014-0155-x)] [Medline: [24777653](https://pubmed.ncbi.nlm.nih.gov/24777653/)]

8. Thackeray R, Neiger BL, Burton SH, Thackeray CR. Analysis of the purpose of state health departments' tweets: information sharing, engagement, and action. *J Med Internet Res* 2013;15(11):e255 [FREE Full text] [doi: [10.2196/jmir.3002](https://doi.org/10.2196/jmir.3002)] [Medline: [24217361](https://pubmed.ncbi.nlm.nih.gov/24217361/)]
9. Harris JK, Moreland-Russell S, Tabak RG, Ruhr LR, Maier RC. Communication about childhood obesity on Twitter. *Am J Public Health* 2014 Jul;104(7):e62-e69 [FREE Full text] [doi: [10.2105/AJPH.2013.301860](https://doi.org/10.2105/AJPH.2013.301860)] [Medline: [24832138](https://pubmed.ncbi.nlm.nih.gov/24832138/)]
10. Duke JC, Hansen H, Kim AE, Curry L, Allen J. The use of social media by state tobacco control programs to promote smoking cessation: a cross-sectional study. *J Med Internet Res* 2014;16(7):e169 [FREE Full text] [doi: [10.2196/jmir.3430](https://doi.org/10.2196/jmir.3430)] [Medline: [25014311](https://pubmed.ncbi.nlm.nih.gov/25014311/)]
11. Chretien KC, Farnan JM, Greysen SR, Kind T. To friend or not to friend? Social networking and faculty perceptions of online professionalism. *Acad Med* 2011 Dec;86(12):1545-1550. [doi: [10.1097/ACM.0b013e3182356128](https://doi.org/10.1097/ACM.0b013e3182356128)] [Medline: [22030752](https://pubmed.ncbi.nlm.nih.gov/22030752/)]
12. DeCamp M, Cunningham AM. Social media: the way forward or a waste of time for physicians? *J R Coll Physicians Edinb* 2013;43(4):318-322. [doi: [10.4997/JRCPE.2013.411](https://doi.org/10.4997/JRCPE.2013.411)] [Medline: [24350316](https://pubmed.ncbi.nlm.nih.gov/24350316/)]
13. Yom-Tov E, Borsa D, Cox IJ, McKendry RA. Detecting disease outbreaks in mass gatherings using Internet data. *J Med Internet Res* 2014;16(6):e154 [FREE Full text] [doi: [10.2196/jmir.3156](https://doi.org/10.2196/jmir.3156)] [Medline: [24943128](https://pubmed.ncbi.nlm.nih.gov/24943128/)]
14. Harris JK, Mansour R, Choucair B, Olson J, Nissen C, Bhatt J, Centers for Disease Control and Prevention. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. *MMWR Morb Mortal Wkly Rep* 2014 Aug 15;63(32):681-685 [FREE Full text] [Medline: [25121710](https://pubmed.ncbi.nlm.nih.gov/25121710/)]
15. Yuan P, Bare MG, Johnson MO, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. *J Med Internet Res* 2014;16(5):e117 [FREE Full text] [doi: [10.2196/jmir.3229](https://doi.org/10.2196/jmir.3229)] [Medline: [24784982](https://pubmed.ncbi.nlm.nih.gov/24784982/)]
16. DeCamp M, Koenig TW, Chisolm MS. Social media and physicians' online identity crisis. *JAMA* 2013 Aug 14;310(6):581-582 [FREE Full text] [doi: [10.1001/jama.2013.8238](https://doi.org/10.1001/jama.2013.8238)] [Medline: [23942675](https://pubmed.ncbi.nlm.nih.gov/23942675/)]
17. Crotty BH, Mostaghimi A, Arora VM. Online identities of physicians. *JAMA* 2013 Dec 18;310(23):2566-2567. [doi: [10.1001/jama.2013.282144](https://doi.org/10.1001/jama.2013.282144)] [Medline: [24346998](https://pubmed.ncbi.nlm.nih.gov/24346998/)]
18. Lee JL, Wu AW. Online medical professionalism. *Ann Intern Med* 2013 Jul 16;159(2):158. [doi: [10.7326/0003-4819-159-2-201307160-00020](https://doi.org/10.7326/0003-4819-159-2-201307160-00020)] [Medline: [23856691](https://pubmed.ncbi.nlm.nih.gov/23856691/)]
19. OrganizedWisdom. About Us page. Saving Doctors Time While Creating Healthier, Happier Patients URL: http://www.organizedwisdom.com/about_us/an/med [accessed 2014-10-01] [WebCite Cache ID 6T08ud4M0]
20. Sugawara Y, Narimatsu H, Hozawa A, Shao L, Otani K, Fukao A. Cancer patients on Twitter: a novel patient community on social media. *BMC Res Notes* 2012;5:699 [FREE Full text] [doi: [10.1186/1756-0500-5-699](https://doi.org/10.1186/1756-0500-5-699)] [Medline: [23270426](https://pubmed.ncbi.nlm.nih.gov/23270426/)]
21. Chretien KC, Azar J, Kind T. Physicians on Twitter. *JAMA* 2011 Feb 9;305(6):566-568. [doi: [10.1001/jama.2011.68](https://doi.org/10.1001/jama.2011.68)] [Medline: [21304081](https://pubmed.ncbi.nlm.nih.gov/21304081/)]
22. Ye Y. Correlates of consumer trust in online health information: findings from the health information national trends survey. *J Health Commun* 2011 Jan;16(1):34-49. [doi: [10.1080/10810730.2010.529491](https://doi.org/10.1080/10810730.2010.529491)] [Medline: [21086209](https://pubmed.ncbi.nlm.nih.gov/21086209/)]
23. Guseh JS, Brendel RW, Brendel DH. Medical professionalism in the age of online social networking. *J Med Ethics* 2009 Sep;35(9):584-586. [doi: [10.1136/jme.2009.029231](https://doi.org/10.1136/jme.2009.029231)] [Medline: [19717700](https://pubmed.ncbi.nlm.nih.gov/19717700/)]
24. Batt-Rawden S, Flickinger T, Weiner J, Cheston C, Chisolm M. The role of social media in clinical excellence. *Clin Teach* 2014 Jul;11(4):264-269. [doi: [10.1111/tct.12129](https://doi.org/10.1111/tct.12129)] [Medline: [24917094](https://pubmed.ncbi.nlm.nih.gov/24917094/)]
25. Decamp M. Physicians, social media, and conflict of interest. *J Gen Intern Med* 2013 Feb;28(2):299-303 [FREE Full text] [doi: [10.1007/s11606-012-2251-x](https://doi.org/10.1007/s11606-012-2251-x)] [Medline: [23129160](https://pubmed.ncbi.nlm.nih.gov/23129160/)]
26. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One* 2010;5(11):e14118 [FREE Full text] [doi: [10.1371/journal.pone.0014118](https://doi.org/10.1371/journal.pone.0014118)] [Medline: [21124761](https://pubmed.ncbi.nlm.nih.gov/21124761/)]
27. Nsoesie E, Marathe M, Brownstein J. Forecasting peaks of seasonal influenza epidemics. *PLoS Curr* 2013;5:- [FREE Full text] [doi: [10.1371/currents.outbreaks.bb1e879a23137022ea79a8c508b030bc](https://doi.org/10.1371/currents.outbreaks.bb1e879a23137022ea79a8c508b030bc)] [Medline: [23873050](https://pubmed.ncbi.nlm.nih.gov/23873050/)]
28. Mishori R, Singh LO, Levy B, Newport C. Mapping physician Twitter networks: describing how they work as a first step in understanding connectivity, information flow, and message diffusion. *J Med Internet Res* 2014;16(4):e107 [FREE Full text] [doi: [10.2196/jmir.3006](https://doi.org/10.2196/jmir.3006)] [Medline: [24733146](https://pubmed.ncbi.nlm.nih.gov/24733146/)]
29. Desai T, Patwardhan M, Coore H. Factors that contribute to social media influence within an Internal Medicine Twitter learning community. *F1000Res* 2014;3:120 [FREE Full text] [doi: [10.12688/f1000research.4283.1](https://doi.org/10.12688/f1000research.4283.1)] [Medline: [25110581](https://pubmed.ncbi.nlm.nih.gov/25110581/)]
30. Himelboim I, Han JY. Cancer talk on twitter: community structure and information sources in breast and prostate cancer social networks. *J Health Commun* 2014;19(2):210-225. [doi: [10.1080/10810730.2013.811321](https://doi.org/10.1080/10810730.2013.811321)] [Medline: [24111482](https://pubmed.ncbi.nlm.nih.gov/24111482/)]
31. Tsuya A, Sugawara Y, Tanaka A, Narimatsu H. Do cancer patients tweet? Examining the twitter use of cancer patients in Japan. *J Med Internet Res* 2014;16(5):e137 [FREE Full text] [doi: [10.2196/jmir.3298](https://doi.org/10.2196/jmir.3298)] [Medline: [24867458](https://pubmed.ncbi.nlm.nih.gov/24867458/)]
32. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: [10.1136/bmj.39489.470347.AD](https://doi.org/10.1136/bmj.39489.470347.AD)] [Medline: [18436948](https://pubmed.ncbi.nlm.nih.gov/18436948/)]

Edited by G Eysenbach; submitted 07.08.14; peer-reviewed by A Dijuricich, H Narimatsu; comments to author 10.09.14; revised version received 12.09.14; accepted 16.09.14; published 15.10.14.

Please cite as:

Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD

What Are Health-Related Users Tweeting? A Qualitative Content Analysis of Health-Related Users and Their Messages on Twitter
J Med Internet Res 2014;16(10):e237

URL: <http://www.jmir.org/2014/10/e237/>

doi: [10.2196/jmir.3765](https://doi.org/10.2196/jmir.3765)

PMID: [25591063](https://pubmed.ncbi.nlm.nih.gov/25591063/)

©Joy L Lee, Matthew DeCamp, Mark Dredze, Margaret S Chisolm, Zackary D Berger. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 15.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Case Study of the New York City 2012-2013 Influenza Season With Daily Geocoded Twitter Data From Temporal and Spatiotemporal Perspectives

Ruchit Nagar^{1,2}; Qingyu Yuan^{1,3}, PhD; Clark C Freifeld⁴, PhD; Mauricio Santillana^{1,5,6}, PhD; Aaron Nojima^{1,7}; Rumi Chunara^{1,8}, PhD; John S Brownstein^{1,6,8}, PhD

¹Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA, United States

²Yale University, New Haven, CT, United States

³Management School, University of Chinese Academy of Sciences, Beijing, China

⁴Boston University, Biomedical Engineering Department, Boston, MA, United States

⁵Harvard University, School of Engineering and Applied Sciences, Cambridge, MA, United States

⁶Harvard University, School of Public Health, Boston, MA, United States

⁷Massachusetts Institute of Technology, Cambridge, MA, United States

⁸Department of Pediatrics, Harvard Medical School, Boston, MA, United States

Corresponding Author:

Ruchit Nagar

Children's Hospital Informatics Program

Boston Children's Hospital

1 Autumn Street

Boston, MA,

United States

Phone: 1 2817258062

Fax: 1 6177300267

Email: ruchit.nagar@yale.edu

Abstract

Background: Twitter has shown some usefulness in predicting influenza cases on a weekly basis in multiple countries and on different geographic scales. Recently, Broniatowski and colleagues suggested Twitter's relevance at the city-level for New York City. Here, we look to dive deeper into the case of New York City by analyzing daily Twitter data from temporal and spatiotemporal perspectives. Also, through manual coding of all tweets, we look to gain qualitative insights that can help direct future automated searches.

Objective: The intent of the study was first to validate the temporal predictive strength of daily Twitter data for influenza-like illness emergency department (ILI-ED) visits during the New York City 2012-2013 influenza season against other available and established datasets (Google search query, or GSQ), and second, to examine the spatial distribution and the spread of geocoded tweets as proxies for potential cases.

Methods: From the Twitter Streaming API, 2972 tweets were collected in the New York City region matching the keywords "flu", "influenza", "gripe", and "high fever". The tweets were categorized according to the scheme developed by Lamb et al. A new fourth category was added as an evaluator guess for the probability of the subject(s) being sick to account for strength of confidence in the validity of the statement. Temporal correlations were made for tweets against daily ILI-ED visits and daily GSQ volume. The best models were used for linear regression for forecasting ILI visits. A weighted, retrospective Poisson model with SaTScan software (n=1484), and vector map were used for spatiotemporal analysis.

Results: Infection-related tweets ($R=.763$) correlated better than GSQ time series ($R=.683$) for the same keywords and had a lower mean average percent error (8.4 vs 11.8) for ILI-ED visit prediction in January, the most volatile month of flu. SaTScan identified primary outbreak cluster of high-probability infection tweets with a 2.74 relative risk ratio compared to medium-probability infection tweets at $P=.001$ in Northern Brooklyn, in a radius that includes Barclay's Center and the Atlantic Avenue Terminal.

Conclusions: While others have looked at weekly regional tweets, this study is the first to stress test Twitter for daily city-level data for New York City. Extraction of personal testimonies of infection-related tweets suggests Twitter's strength both qualitatively and quantitatively for ILI-ED prediction compared to alternative daily datasets mixed with awareness-based data such as GSQ. Additionally, granular Twitter data provide important spatiotemporal insights. A tweet vector-map may be useful for visualization of city-level spread when local gold standard data are otherwise unavailable.

(*J Med Internet Res* 2014;16(10):e236) doi:[10.2196/jmir.3416](https://doi.org/10.2196/jmir.3416)

KEYWORDS

influenza; Twitter; New York City; spatiotemporal; Google Flu Trends; infodemiology; mHealth; social media, natural language processing; medical informatics

Introduction

Seasonal influenza epidemics result in an estimated three to five million cases of severe illness and 250,000 to 500,000 deaths worldwide each year [1]. In order to better control seasonal influenza epidemics and the next pandemics, researchers have proposed several “infodemiological” approaches [2,3] to achieve near real-time surveillance with Internet data, such as Google search query (GSQ) [4-8] and textual Twitter data [9-15].

The idea of using GSQ data volume for detecting outbreaks was first introduced in 2006 by Eysenbach et al [4], then refined by Ginsburg et al with the Google Flu Trends project [5]. Until recently, Google Flu Trends has matched weekly influenza-like illness (ILI) incidence well, but the recent 2012-2013 flu season in the United States was overestimated by 200% [16]. Though providing impetus for increased vigilance during one of the worst flu seasons to date, this case demonstrates that Google Flu Trends alone has limitations for influenza prediction, and motivates search for improved models [17,18]. New models can look for refined algorithms but may also seek out independent datasets to improve predictive abilities. With personal, textual data from potentially sick tweeters, Twitter offers a rich alternative dataset. Twitter epidemic surveillance was first introduced by Ritterman et al (2009) [9], who showed that monitoring tweets can improve the accuracy of market forecasting models by providing early warnings of external events like the H1N1 outbreak. Chew and Eysenbach subsequently evaluated the content and sources of the Twitter response to the swine flu outbreak in 2009, seeing potential for public health insights [19]. Recent papers from Culotta [20,21], Signorini [22], Vadileios [10], Kim [12], Santos [23], and Achrekar [15] et al have shown the usefulness of Twitter data to retrospectively predict national and regional influenza cases on a weekly basis. Broniatowski and colleagues have recently considered Twitter's temporal predictive strength for cases at the municipal level in New York City, with a prospective approach [24].

The promise of Twitter as a mechanism for timely signal detection has spawned efforts to build tools for surveillance, including MappyHealth [25], germTracker [26], Observatorio da Dengue [27], Infovigil [28], and SickWeather [29]. But studies of Twitter's temporal predictive capabilities have not adequately addressed how the information can be brought to a concrete, local scale. In this paper, we look to dive deeper into the case of New York City (NYC) by analyzing daily Twitter

data from temporal and spatiotemporal models. We look for novel insights by being the first to consider daily data (Twitter and GSQ) for influenza case predictions and by using geolocated tweets to estimate the probability of flu transmission within the city.

Methods

Data Sources

Official influenza-like illness emergency department visit (ILI-ED) counts were obtained from the Weekly Influenza Surveillance Reports released by the municipal government of the city of New York [30]. These reports were compiled into a daily breakdown. The data for the ILI cases was extracted from the graphs using optical plot reading software (WebPlotDigitizer [31]) validated by pixel counting for each daily value. Data for total counts for NYC and counts for each borough were determined by this method. This method was employed because data were otherwise unavailable from the NYC Department of Health.

Google Trends Search Query Data

Google Trends provides a time series index of the volume of queries users enter into Google in a given geographic area. The query index is based on query share: the total query volume for the search term in question within a particular geographic region divided by the total number of queries in that region during the time period being examined. The maximum query is normalized to 100 and the query share at the initial date being examined is normalized to be zero [7]. This query share also varies with time. In our model, the search queries used were “flu”, “gripe”, “influenza”, and “high fever”. According to our survey, these queries gave the highest signal. More importantly, these exact keywords were also used for tweet collection as a basis for comparison. Google provides weekly data for state and (limited) city searches. Daily data had to be extracted by querying the search volume over overlapping time periods and proportionally adjusting the query index along the time series. The Trends data were all downloaded within a single day (May 11, 2013), as Google varies the signal display with time. It is important to note that these data are distinct from the weekly Google Flu Trends (GFT) data, which does not provide a daily breakdown and was not used in this study. That being said, GSQ could act as a potential proxy for GFT. When comparing the 7-day GSQ volume totals with GFT weekly volume hits, their correlation coefficient was .78 between September 23, 2012 and May 5, 2013.

Twitter Data

Using the Twitter Streaming API, geocoded tweets were collected from October 15, 2012 to May 10, 2013; unlike the method suggested by Broniatowski and colleagues, our approach did not involve normalizing the flu-related tweets against the weekly or daily total Tweet count. This was not done because we saw variation in the baseline number of tweets per day and because the daily volume of tweets was low (ranging from 0 to 120 tweets at peak season).

These tweets were selected based on the geographical bounding box from (40.44, -74.93) to (41.12, -72.63). This window was chosen to account for people commuting into New York City from New Jersey and Long Island. The assumption was that the signal position of tweets would not change significantly by moving the bounding box further out from the city; however, the larger area would allow for greater tweet volume for analysis. To determine if the tweet was inside the frame of interest, the tweet latitude and longitude were first checked. If the tweet location was missing, the profile latitude and longitude was used. Users could also define a text-based location for their profile, but these tweets could not be reliably determined to be inside the bounding box, and were therefore excluded. Dredze et al have developed “Carmen”, a system to geolocate tweets by cross-referencing location keywords to a database, but there remains difficulty in comparing tweets with a GPS stamp to those which are geolocated to a broader region [32].

Keyword filters were then applied to our collection of tweets: “flu”, “gripe”, “influenza”, and “high fever” were case-insensitive, word-bounded inclusion strings. “Avian”, “stomach”, and “bird” were exclusion strings. Although the purpose of this study was not to optimize keyword selection, preliminary studies were conducted to see the effect on the signal to noise ratio when adding inclusion words for other ILI symptoms and medications. Ginsberg and colleagues methodically constructed a set of 45 significant keywords for Google Flu Trends using a linear regression model [5]. Similar methods have been used by Kim et al for Hangeul Twitter [12]. Optimal keywords, however, vary across time and geographical region, so our approach focused on obtaining a signal that was both strong and specific and was not concerned with the decreasing marginal value of the next best keyword. Our search was broad enough to cover the top 11 ILI-related search query topics that grouped the Google Flu Trends 45 keywords.

After collecting the filtered tweets, all 2972 tweets were manually curated. Duplicate tweets from the same user were first removed from the dataset. We then created categories for Twitter classification using the models established by Lamb et al [33]: Relevant (R) vs Irrelevant (Ir), Awareness (A) vs Infection (I), and Self (S) vs Other (O). Additionally, a fourth option (not included by Lamb et al) was added to mark High (H) vs Medium (M) vs Low (L)—a guess for the probability of the subject(s) mentioned being sick. This category could help differentiate a sarcastic tweet from a serious one, by assigning a rank to the veracity of the subject being sick. A user-based guess for the sickness of the subject has also been employed by germTracker and CrowdBreaks to leverage human-based classification of tweets [34].

All tweets were labeled after being dissociated with their dates to prevent bias of an anticipated regular flu season on categorization. The categories were grouped into 12 four-letter codes and an additional code for irrelevant tweets. In general, Relevant-Infection-Self-High/ Relevant-Infection-Self-Medium (RISH/RISM) and Relevant-Infection-Other-High/ Relevant-Infection-Other-Medium (RIOH/RIOM) constituted tweets for people who had the flu or flu-like symptoms; Relevant-Infection-Self-Low/ Relevant-Infection-Other-Low (RISL/RIOL) grouped tweets for people who were recovering from illness; Relevant-Awareness-Self-High/ Relevant-Awareness-Self-Medium (RASH/RASM) generally grouped tweets of people with negative reactions to flu shots; Relevant-Awareness-Self-Low (RASL) grouped tweets for flu shots and successful vaccinations; Relevant-Awareness-Other-High / Relevant-Awareness-Other-Medium (RAOH/RAOM) referred to news media alerts of different severity; and Relevant-Awareness-Other-Low (RAOL) grouped tweets focused on public health awareness and therapies. Examples of this scheme can be seen in [Multimedia Appendix 1](#).

Ethics

The tweets used in this study were publicly distributed. Consent for these tweets to be read comes from users signing the Twitter Terms and Agreement and agreeing to public privacy settings. This project was exempted from IRB approval since it did not meet the criteria for human subjects’ research. Nevertheless, for the purpose of this study, tweeters’ user IDs were not collected so each tweet entity remained anonymized. The subsequent analyses were dependent on the tweet content, tweet frequency (aggregate), and tweet location.

Predictive Models

Temporal

First, Pearson correlation values were constructed between each of the Twitter category time series and the GSQ time series to the ILI-ED time series for New York City. Next, to forecast the ILI data, AR (auto-regressive) models were used. ILI was the dependent variable, while GSQ and the strongest category Twitter data were the independent variables separately in different models. Each model was tested for 7 weeks between January 6, 2013 and February 23, 2013 to compare the predictive abilities in ILI visits during the volatile peak of the flu season.

Spatiotemporal (Retrospective)

Since the selected tweets are geocoded and include a date stamp, they are also fact space-time data. Using SaTScan software [35], we were able to perform retrospective geographical surveillance of the 2012-2013 flu season to detect statistically significant space-time disease clusters. The space-time test statistic is defined by a cylindrical window with circular base. The circular base represents the spatial scan, with the base varying from zero to a size that captures 50% of the population risk for a given tweet. The set of circular bases scan the map for potential clusters. Similarly, the height of the cylinder represents the temporal progression of the map varying from 0 to 50% of the total time period. The cylinder is moved both laterally in space and vertically in time (aggregated in weekly steps) to identify

possible clusters for the entire study region to generate the test statistic.

A total of 1484 total RISM and RISH tweets, restricted to New York City proper, were used for the spatiotemporal analysis between October 15, 2012 and May 10, 2013. RISH tweets were treated as the potential flu cases, while RISM tweets were treated as control cases. These tweets were also weighted based on their seasonal time-series correlation value to gold standard ILI visits (RISH=.689, RISM=.655) to better reflect their relative association with real flu cases. After assigning weights, a Poisson probability model was used to search for high clusters with an excess of RISH tweets (compared to control RISM tweets), for each cylinder in the scan window. Both primary and secondary clusters were identified based on a likelihood ratio test statistic using the methods described by Jung and Kulldroff [36].

Spatiotemporal (Prospective)

SaTScan provides prospective analysis functionality to predict regions with higher risk ratio. We looked at representing these results in an alternative fashion, in a manner accessible from a Web application. The approach was to construct a vector-map for the larger NYC region. The idea was to use an analogy of a weather map where winds move from areas of high to low pressure. Pressure was modeled by density of potentially sick tweeters and winds represented the direction of change locally in that density. Specifically, for each 0.1 x 0.1 decimal degree grid, the weekly change in percent of flu-related tweets was calculated. A vector was drawn to point in the direction of the neighboring cells with the highest positive percentage increase in flu-related tweets, away from cells with a percentage decrease. Red corresponds with higher percent increase. The underlying assumption of the model was that infection spreads locally between neighboring regions. While it may be argued that disease transmission is not spatially continuous, this model demonstrates one of many possible representations of flu dynamics at the city-level scale (hence the grid dimensions),

and how a real-time public health tool may be displayed. Sadilek et al have suggested the importance of colocation with other sick tweeters (friends or otherwise) in their individual-based model of spatiotemporal prediction. Therefore, understanding changes in sick tweeter colocation is importantly indicated by our wind-map. The macro-level validation of our model will come only with the availability of more gold-standard, spatiotemporal data.

Further discussion of spatial models of temporal windows and their relation to spatial predictor covariates (age, ethnicity, population density, distance to school and subway, distance from home, distance from vaccination sites) can be found in Figures 8 and 9 in [Multimedia Appendix 2](#).

Results

Temporal

The time series of tweet counts was calculated to first assess the quality of the data. We classified over 90% of the tweets collected as relevant. Most tweets were about the self (S), about being infected (I) or having ILI symptoms, and were ranked with high probability (H) for the subject being sick, as seen in [Table 1](#). These Pearson correlation values correspond to tweets from October 15, 2012 to May 10, 2013.

Each time series, including the GSQ time series, was then compared to the ILI data and ranked (as seen in [Table 1](#)). Infection, RISH, and Relevant groups and subgroup had higher Pearson correlations with the gold-standard ILI data between October 15, 2012 and May 10, 2013. [Figure 1](#) illustrates aforementioned relationships with the ILI data.

[Figure 2](#) illustrates the comparison between GSQ and Twitter data, comparing both to ILI. Note the Pearson correlation between GSQ and Twitter "Awareness" time series is .934. Both have a characteristic spike on January 10, 2013, the day the city of Boston declared a public health emergency for influenza.

Table 1. Quality of the classified tweets and search query data.

Tweet group ^a	Percentage of tweets	Time series	Pearson correlation
Relevant	0.907	Infection	.763
Self	0.689	RISH	.689
Infection	0.628	Relevant	.687
High	0.497	GSQ	.683
Awareness	0.279	Self	.677
Medium	0.223	Medium	.668
Other	0.219	Other	.666
Low	0.188	High	.665
Irrelevant	0.082	RISM	.655
		RIOH	.616
Sub Group		RAOM	.587
	RISH	Awareness	.549
	RASL	RASM	.545
	RISM	RISL	.542
	RAOM	RIOM	.511
	RIOH	Low	.451
	RISL	RAOH	.411
	RAOH	RASL	.351
	RASM	RASH	.322
	RAOL	RAOL	.277
	RIOM	RIOL	.254
	RIOL	Irrelevant	.213
	RASH		

^aRelevant (R), Awareness (A), Infection (I), Self (S), Other (O), High (H), Medium (M), Low (L).

Figure 1. Time series comparisons between Tweet categories and ILI-ED visits.

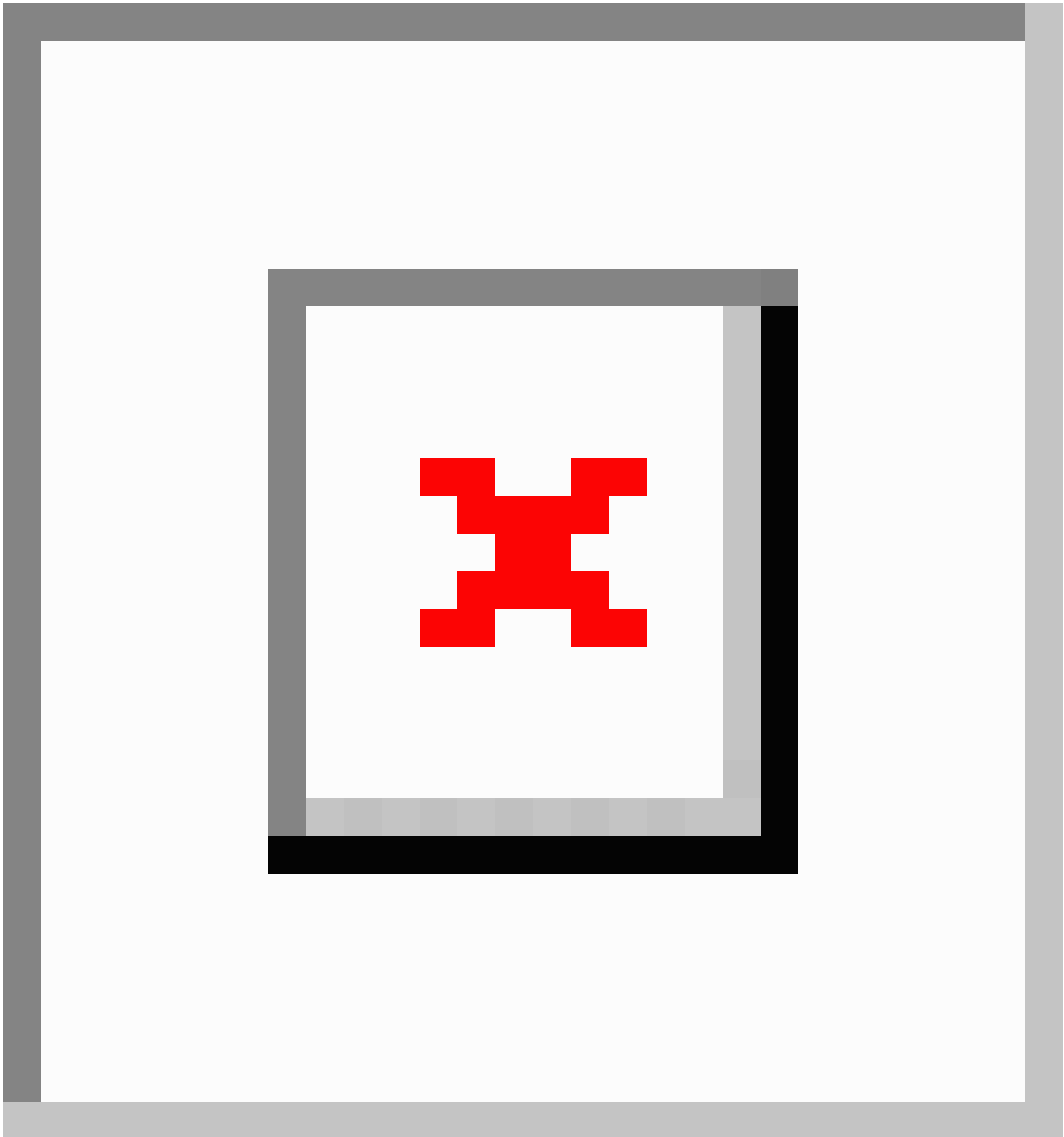
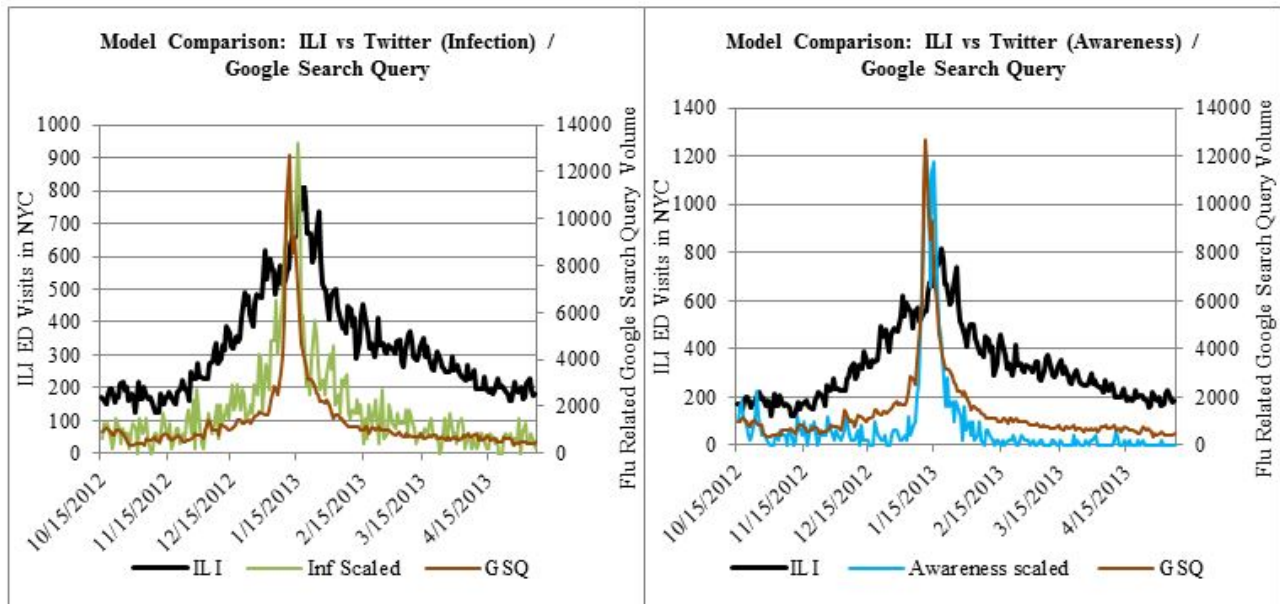


Figure 2. Comparison of Infection tweets and Awareness-based data.



Linear Regression Models

In order to compare the Twitter, GSQ, and ILI time series datasets, we first used an Augmented Dickey-Fuller (ADF) test with EVIEWS-7 statistical package to confirm the presence of a unit root. Each dataset was then adjusted to meet the stationarity test. The datasets were not stationary at the 1% significance level until after first order differencing was performed (Tables 2 and 3).

An Englemen Granger co-integration test was then performed on the time-lagged datasets. Establishing co-integration and stationarity allowed for determination of consistent estimators in our regression model. One weakness was present in our approach: seasonality was not tested because the available data only comprised one season. A discussion of the weekday effect and why weekly cyclic terms were not included in the model can be found in Multimedia Appendix 2.

The first AR model incorporates infection time series data and 14-day time-lagged ILI data for the first week. The time lag reflects the real delay in publicly reporting updated ILI-ED visit counts. The model and results are shown in Table 4. Note mu and epsilon are terms for error and drift respectively in the subsequent models.

$$\text{Model 1: } ILI = m_1 ILI(-14) + m_2 \text{infection}(-1) + m_3 \mu(-1) + \epsilon$$

The results from the model above show that the time-lagged ILI data were not significant for the prediction. This persuaded us to modify the model to incorporate only Infection tweets

time series data. Across all 7 weeks, the best model for the Infection tweets time series had the following form:

$$\text{Model 2: } ILI = \alpha_1 \text{infection}(-2) + \alpha_2 \mu(-1) + \alpha_3 \mu(-2) + \epsilon$$

We repeated the same process for GSQ data, constructing a model to incorporate both it and ILI time-lagged data for the first week. The results are listed in Table 5.

$$\text{Model 3: } ILI = \beta_1 ILI(-14) + \beta_2 \text{gsq}(-3) + \beta_3 \mu(-1) + \beta_4 \mu(-2) + \epsilon$$

Again we found the time-lagged ILI data to be not significant and eliminated it from the subsequent GSQ models. For the Google search query data, the following models were constructed to optimize its predictive scores by adjusting the time series lag.

$$\text{Model 4: } ILI = \gamma_1 \text{gsq}(-3) + \gamma_2 \mu(-1) + \gamma_3 \mu(-2) + \epsilon \text{ (first and second weeks)}$$

$$\text{Model 5: } ILI = \gamma_1 \text{gsq}(-4) + \gamma_2 \mu(-1) + \gamma_3 \mu(-2) + \epsilon \text{ (third to seventh weeks)}$$

The two GSQ and the Twitter Infection models were then compared by their mean absolute percent error or MAPE (see Table 6). The data suggest lower MAPE scores for Twitter for the first 4 weeks in January (4.7, 6.9, 11.8, and 10.4) compared to GSQ data (5.5, 15.8, 12.4, and 11.3).

The average MAPE for temporal predictions using the Infection tweet time series was 8.4. Figure 3 demonstrates the ILI predictions using the Infection tweet time series (model 2) for the month of January.

Table 2. Augmented Dickey-Fuller (ADF) test of ILI^a, Twitter, and Google search query data.

	ILI		Twitter infection		Google search query	
	<i>t</i> statistic ^b	probability	<i>t</i> statistic ^b	probability	<i>t</i> statistic ^b	probability
ADF test	-1.902	0.331	-2.569	0.101	-2.844	0.054
Test critical values						
1% level	-3.462		-3.463		-3.463	
5% level	-2.876		-2.876		-2.876	
10% level	-2.574		-2.574		-2.574	
	Non-stationary		Non-stationary		Non-stationary	

^aILI: influenza-like illness

^bDegrees of freedom=203

Table 3. Augmented Dickey-Fuller (ADF) test of ILI^a, Twitter, and Google search query data with first order lag.

	Δ ILI ^b		Δ Twitter infection ^b		Δ Google search query ^b	
	<i>t</i> statistic ^c	probability	<i>t</i> statistic ^c	probability	<i>t</i> statistic ^c	probability
ADF test	-12.544	0.000	-19.358	0.000	-6.920	0.000
Test critical values						
1% level	-3.463		-3.463		-3.463	
5% level	-2.876		-2.876		-2.876	
10% level	-2.574		-2.574		-2.574	
	Stationary		Stationary		Stationary	

^aILI: influenza-like illness

^b Δ =first order lag

^cDegrees of freedom=202

Table 4. Results of model (1).

Variable	Coefficient	Standard error	<i>t</i> statistic ^c	Probability
Infection(-1)	-2.174	1.016	-2.140	0.036
ILI ^a (-14)	0.224	0.142	1.576	0.120
AR ^b (1)	1.007	0.016	61.676	0.000

^aILI: influenza-like illness

^bAR: auto-regressive

^cDegrees of freedom=188

Table 5. Results of model (3).

Variable	Coefficient	Standard error	<i>t</i> statistic ^d	Probability
GSQ ^a (-3)	0.069	0.031	2.218	0.030
ILI ^b (-14)	0.212	0.147	1.444	0.154
AR ^c (1)	0.690	0.125	5.515	0.000
AR(2)	0.315	0.127	2.476	0.016

^aGSQ: Google Trends search query

^bILI: influenza-like illness

^cAR: auto-regressive

^dDegrees of freedom=188

Table 6. MAPE^a scores for Infection tweet and GSQ^b models for ILI^c predictions.

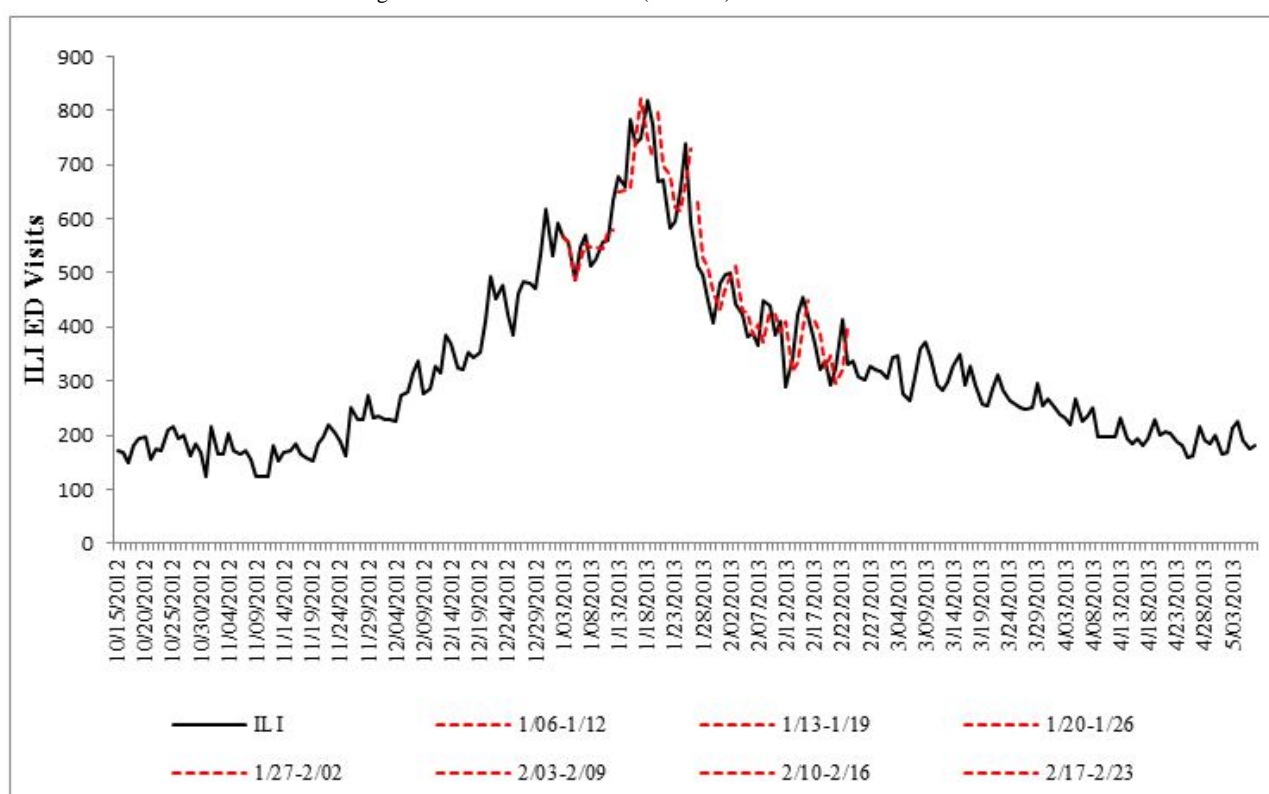
Date	Twitter model		GSQ models	
	Durbin-Watson statistic	MAPE (static)	Durbin-Watson statistic	MAPE (static)
1/06-1/12	2.00	4.7	2.04	5.5
1/13-1/19	2.11	6.9	2.13	15.8
1/20-1/26	2.16	11.8	2.16	12.4
1/27-2/02	2.07	10.4	2.04	11.3
2/03-2/09	2.09	8.2	2.06	7.9
2/10-2/16	2.08	14.8	2.05	15.2
2/17-2/23	2.08	15.3	2.05	14.5

^aMAPE: mean absolute percent error

^bGSQ: Google Trends search query

^cILI: influenza-like illness

Figure 3. Predicted ILI-ED visits in red using the Infection tweets model (Model 2).

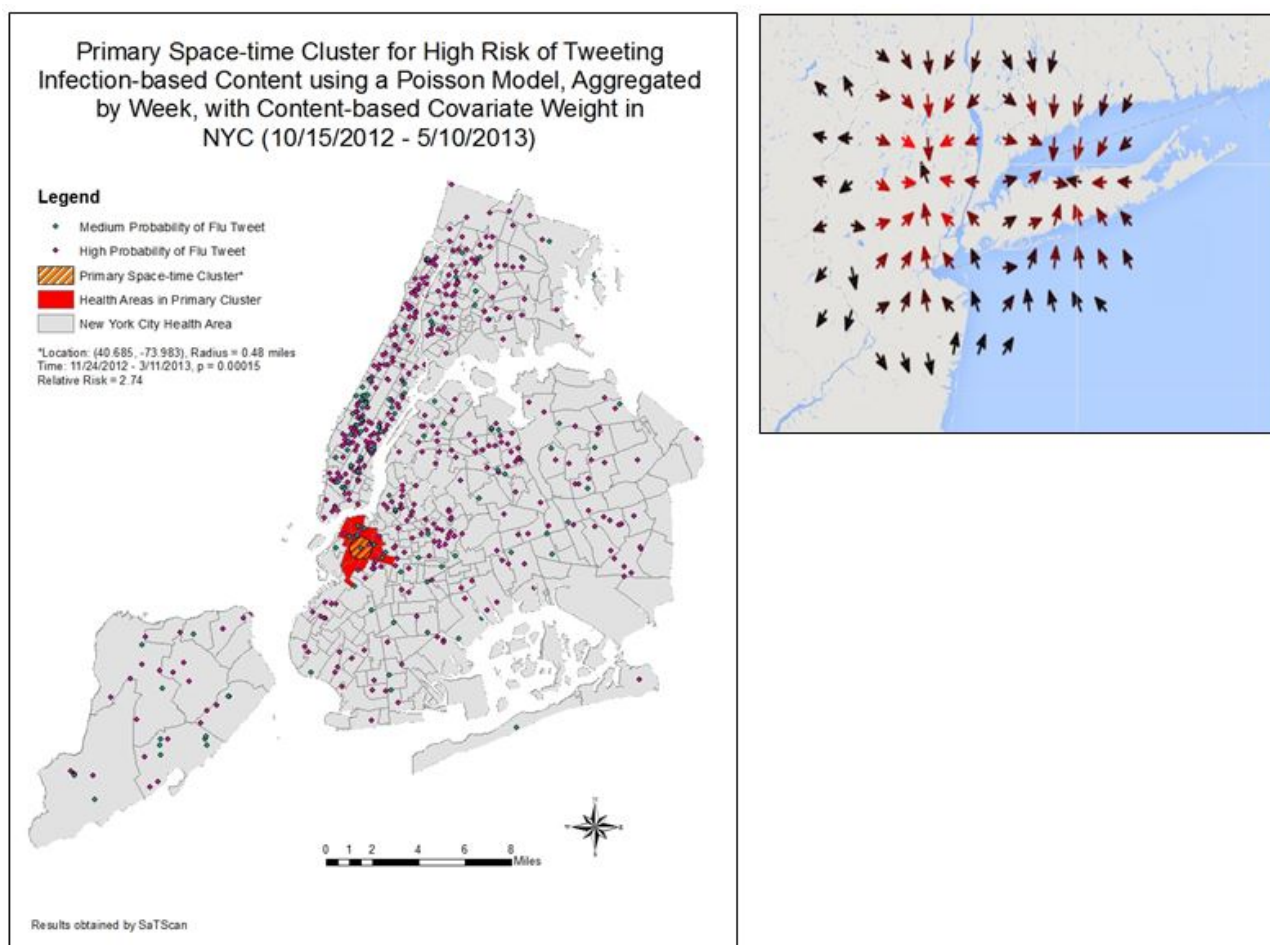


Spatiotemporal Models

From the retrospective analysis [37], a primary space-time cluster for RISH tweets was found in North Brooklyn [38] with a relative risk of 2.74 (RISH to control RISM) between

November 24, 2012 and March 11, 2013 at a significance of $P < .001$. The prospective vector-map was constructed for the week following January 8-15, 2013 and can be seen in Figure 4.

Figure 4. Right: Retrospective primary space-time cluster ($p < .001$) for high risk of tweeting flu infection-based Content, determined by a Poisson Model with cases as High Probability Flu Tweets and controls as Medium Probability Flu Tweets, aggregated by week, and with content-specific covariate weight in NYC during 10/15/2012-5/10/2013. Top Left: Epicenter located at (40.685, -79.983) with 0.48 mile radius including places of mass gathering such as Barclays Center and Atlantic Avenue Terminal. Bottom Left: Prospective approach to modeling weekly changes in local Infection-tweet spread.



Discussion

Principal Results

The principal aim of this study was to evaluate the strength of Twitter data in predicting flu cases on a daily local level in New York City. Our method for selecting this Twitter data with keyword filters and location filters returned a dataset with 90%+ relevant tweets. Daily Infection tweets showed the highest correlation with daily ILI visits ($R=.763$) for the 2012-2013 time period. While we believed that the RISH category most clearly identified personal accounts of sickness, the correlation with ILI was not as strong as using just the Infection tweets umbrella ($R=.689$). This may be due to the fact that Infection aggregates reports from both individuals and others to better reflect total cases of flu across various levels of certainty based on level of symptom progression. Low correlation values for the RISL group (.542) compared to the correlations for RISM and RIOH (.616 and .655), however, still are consistent with the evaluator guess in the classification scheme.

Twitter data (Infection and RISH) importantly outperformed daily GSQ data ($R=.683$) with respect to correlation to daily ILI visits between October 15, 2012 to May 10, 2013. Qualitatively these results are consistent with our expectations. Twitter allows

infection-related testimonials to be extracted. These Infection tweets are better predictors of potential cases than Awareness tweets (Figure 1) and GSQ data, neither of which can distinguish people who search due to awareness or from infection. In fact, GSQ resembles Awareness tweets (Figure 2), with a correlation of .934. Both datasets had large increases (about 600%) in volume immediately after the nearby January 10 Boston public health emergency declarations (the New York state of emergency was announced 2 days later and also showed a spike). Similar spikes in public attention on Twitter after media releases have been noted by Gu and colleagues in their analysis of the 2013 H7N9 outbreak in China, with the most pronounced effect in the first 3 days [39]. From a public health standpoint, an awareness reaction to a media response provides information that is either exaggerated, already established, or both. Infection tweets are more relevant because they indicate current probable cases not necessarily accounted for by the hospital network. This is not to entirely discount Awareness tweets. In fact, Awareness-related tweets can still provide important insight on vaccination. The data from Relevant, Awareness, and Self categories suggests people tweeted about flu shots when it was probably too late (see Figure 2 in Multimedia Appendix 2). Although there was some signal from October to December, tweets about the flu shot peaked near peak flu season, at which

point those tweeters would have already been exposed to the virus with weakened immunity. It is possible that delayed vaccinations contributed to the increased intensity of the peak season.

When considering temporal forecasting, Infection tweets performed better than the GSQ data. While the predictions were comparable for the off-peak flu season, the difference in MAPE was apparent between January 6 and February 23, 2013, which included the peak and highest volatility of the flu season. The Infection tweet model had an average MAPE of 8.4 compared to GSQ models that had a MAPE of 11.8. Importantly, available and officially released time-lagged ILI data were found to be not statistically significant in predicting real-time ILI cases. This further underscores the need for alternative real-time data sources such as Twitter.

Spatiotemporal analysis also provided valuable insights. In particular, a primary cluster of a high ratio of high-probability sick tweeters to medium-probability infection tweeters was found in northern Brooklyn across a timeframe from November through March. This cluster includes Barclay's Center and the Atlantic Avenue Terminal—both places of mass gathering and commute and therefore increased probability of infection and/or transmission of influenza. It is not definitive what factors led to an increased propensity to tweet high-confidence infection tweets or by extension what factors led to increased sickness in that cluster. The New York City Department of Health does have daily time series data for each borough (see [Figure 3](#) in [Multimedia Appendix 2](#)). The Bronx, Manhattan, Brooklyn, and Queens all tend to follow similar trends in the overall ILI-ED visits as the whole of NYC; each of the four boroughs shares the peak flu visits for the period between January 8 and January 18, 2013. Staten Island has far fewer reports and no discernible peak. Aggregating the Department of Health time series data on the borough level may not immediately reveal knowledge of potential clusters, such as the one identified in Northern Brooklyn by SaTScan using geocoded tweets. Prospective models, like the one in this study (see [Figure 4](#)), have yet to be validated due to lack of geographical gold standard data for comparison. We believe that the representation however is useful in demonstrating one possible model of local diffusion.

Limitations

This study faced several limitations. Classifying the tweets underscored that textual interpretation is a difficult task that requires human interpreters with contextual knowledge of the language and region of interest. The ability to recognize slang, misspellings, Twitter lexicon, inside references, current events, intention, and mood of the tweet sets a high threshold for extracting meaning and sentiment for machine-learning algorithms, experienced researchers, and contracted data classifiers alike. For metropolitan areas that have higher tweet density, multiple languages can come into play. The word “gripe” for example can mean complaint in English but influenza in Portuguese and Spanish; it is also a misspelling of “grippe” in French. Moreover, the tweets were only queried in English and Spanish. With already low tweet volume, capturing other languages such as Italian, Portuguese, Malay, and Tagalog will

be needed to refine models moving forward. When it comes to qualitative coding, checking for inter-rater reliability is key as the process is inherently subjective. We are attaching our dataset in [Multimedia Appendix 3](#) to improve the feedback of the classification scheme.

The classification approach used here also was manual and the keyword choice was not optimized through iterative deletion from a large bank of keywords. That being said, the purpose was not to obtain the highest correlative value as optimal keywords vary in time frame and region. While the manual approach used here has advantages of eliminating false positives/negatives that may result from automatic classification (as has been commonly reported [40]), the tradeoff of this approach comes in speed of analysis. Automatic classifiers have been successful and can be trained to include additional search strategies for influenza from this study (see [Table 1](#) in [Multimedia Appendix 2](#)) to increase speed and accuracy. These rule-based heuristics supplement a growing base of classifier findings by Paul et al [33] and Nagel et al [41], which show bag-of-words, URL Web addresses, retweet status, emoticons, and syntactical organization of a tweet as indicators of illness.

Finally, limitations were present in the modeling approaches. For temporal modeling, seasonality factors were not considered in the time series analysis. This was due to lack of retrospective ILI and Twitter data access. Since influenza is cyclical, seasonality concerns are of extreme relevance in predicting the weeks of the influenza peak. For spatial models, geocoded tweets are few in volume, presenting a clear limitation to the power of the analysis. While many tweeters in New York City may also be tweeting they are sick (without a geocode), at the moment, it is not possible to verify that they are indeed tweeting from within New York City with the available data. Geocoded tweets are, however, expected to grow in the coming years and with this increase comes the potential for higher statistical power [42]. Moreover, Carmen and other text-mining approaches are being considered to increase the fidelity of non-GPS, location-based data [32,43]. From our survey, New York City had the highest volume of geocoded tweets (from 2-3% of total tweets) [42], so it is not clear if the models used here will hold for other cities where data are less prevalent. It is encouraging that in this case study the number of tweets found statistically significant spatiotemporal clusters and temporal autocorrelations at the $P=.002$ and $P=.01$ levels respectively based on the sample. But even with increased geolocation of tweets, when it comes to tracking disease within cities, two obstacles remain: how to verify tweet content, and how to account for tweeter mobility to treat Infection tweets as footprints rather than static incidents. Interactions with the disease could result from interactions within familiar networks of people or from commuting across vast environments.

Comparison to Prior Work

This is not the first study to demonstrate the relevance of Twitter in predicting influenza cases. Broniatowski and colleagues' recent examination of Twitter in New York City can be used as a basis of comparison for the temporal analysis [24]. Their algorithm for Infection tweets found a stronger correlation to ILI visits than the approach here ($R=.88$ vs $R=.763$).

Broniatowski et al's simpler keyword filter algorithm, however, has a lower correlation value with ILI visits ($R=.72$). These comparisons, however, are not *ceteris paribus*. In this study, daily, not weekly data are correlated to see for the first time if daily data have sufficient signal at the municipal level. It is also not clear which keywords are being employed in their algorithms and how to compare the forecasting models used without MAPE scores for a given week's predicted ILI counts.

From a spatiotemporal basis, it is important to consider how this work relates to the framework proposed by Sadilek et al [44]. The approach here attempts to simply map geographical risk based on density of tweets in a given region. Our models would suggest that colocation with high density RISH tweets would suggest higher risk of contracting the disease. Sadilek and colleagues have included colocation with people in social networks as a facet of their model. Our retrospective and prospective models lacked a covariate to measure this network interaction, and thus leave room for improvement. At the same time, the limited number of publicly available geocoded tweets may suggest that such a framework is difficult to implement at the municipal level for a short time period.

Finally, we see an avenue for improving the classification scheme established by Lamb et al by including an evaluator guess. This factor can account for sarcasm, tone of voice, and confidence of the statement made by the tweeter. This factor was also crucial in establishing a basis for case versus control "sick" tweets for spatiotemporal analysis.

Conclusions

This study presents several major takeaways. It is the first study to consider daily city-level Twitter data as a means of

forecasting real-time ILI emergency department visits in New York City. It also suggests useful models for leveraging the geocoded Twitter data to understand potential hotspots of disease (such as the Barclay's Center and Atlantic Avenue Terminal) as they move throughout the flu season. This information will be relevant in considering availability and access to vaccination sites and monitoring ongoing vaccination rates. Twitter can also inform public health officials of the local, upcoming burden of disease. Public health officials already use SaTScan with electronic medical record (EMR) data to track anomalous outbreaks of disease in space and time. Now, Twitter can provide weighted potential cases from personal reports to improve these models. When hyperlocal, confirmed data on flu cases is otherwise unavailable, Twitter provides a real-time information data source. This information can be filtered to select for infection-specific testimonials and as a dataset, outperforms awareness-mixed, daily data from Google Trends search query. This data can also be leveraged in unique prospective models to forecast ILI trends in space and time (see Figure 5 in [Multimedia Appendix 2](#)).

Moving forward, it will be critical to define the threshold of localization for which Twitter can be a useful predictive dataset. For Twitter data to be validated, gold standard public health data must be made more available. Testing the correlative value for the flu cases at the NYC borough level begins to show the limits of Twitter's capabilities in local ILI-ED prediction (see Figure 7 in [Multimedia Appendix 2](#)). With an expected increase in mobile devices and social media activity in the coming years, we look forward to the new challenges, insights, and applications gained from Twitter in the growing field of data-driven epidemiology.

Acknowledgments

We would like to thank Professors Kristina Talbert-Slagle and Elizabeth Bradley at Yale University for their insight and support during the project. We would also like to acknowledge Twitter for making data available through their API and the Yale Global Health Field Experience Award and NIH Grant: 5 R01 LM010812-05 from the National Library of Medicine, which helped in part to fund this study.

Authors' Contributions

RN, CF, and AN collected data. RN and QY coded the data. RN, QY, and MS analyzed temporal models. RN analyzed spatial and spatiotemporal models. RN wrote the manuscript and made all subsequent revisions. MS, QY, RC, CF, and JSB contributed feedback toward the manuscript. JSB is the primary investigator. The sponsors had no role in the design, analysis, or writing of this manuscript.

Conflicts of Interest

None reported.

Multimedia Appendix 1

Examples of classified tweets.

[[PDF File \(Adobe PDF File\), 205KB - jmir_v16i10e236_app1.pdf](#)]

Multimedia Appendix 2

Search phrases for automated-searching; Consideration of weekday effects in temporal models; spatial models of sick tweeters within NYC and relation to vaccination sites; Temporal trends for Awareness-based data; Vector map construction for online Web application; and Gold standard, daily prediction at NYC borough level.

[[PDF File \(Adobe PDF File\), 843KB - jmir_v16i10e236_app2.pdf](#)]

Multimedia Appendix 3

Flu-related tweets and their associated codes.

[[XLSX File \(Microsoft Excel File\), 690KB - jmir_v16i10e236_app3.xlsx](#)]

References

- World Health Organization. Influenza (Seasonal). 2013. URL: <http://www.who.int/mediacentre/factsheets/fs211/en/> [accessed 2014-03-21] [[WebCite Cache ID 6OFZ1df6y](#)]
- Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. *Am J Prev Med* 2011 May;40(5 Suppl 2):S154-S158. [doi: [10.1016/j.amepre.2011.02.006](https://doi.org/10.1016/j.amepre.2011.02.006)] [Medline: [21521589](#)]
- Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11(1):e11 [[FREE Full text](#)] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](#)]
- Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. In: *AMIA Annu Symp Proc. USA: AMIA; 2006 Presented at: AMIA Annual Symposium; November 11, 2006; Washington, D.C p. 244-248* URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839505/>
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014. [doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634)] [Medline: [19020500](#)]
- Doornik J. University of Oxford, Technical Report: 1-21. Improving the timeliness of data on influenza-like illnesses using Google search data URL: <http://www.gwu.edu/~forcpgm/JurgenDoornik-final-Doornik2009Flu-Jan31.pdf> [accessed 2014-09-28] [[WebCite Cache ID 6SvuZsqBO](#)]
- Choi H, Varian H. Predicting the present with Google Trends. *Economic Record* 2012;88(s1):2-9. [doi: [10.1111/j.1475-4932.2012.00809.x](https://doi.org/10.1111/j.1475-4932.2012.00809.x)]
- Carneiro HA, Mylonakis E. Google Trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis* 2009 Nov 15;49(10):1557-1564 [[FREE Full text](#)] [doi: [10.1086/630200](https://doi.org/10.1086/630200)] [Medline: [19845471](#)]
- Ritterman J, Osborne M, Klein E. Using prediction markets and Twitter to predict a swine flu pandemic. 2009 Presented at: 1st International Workshop on Mining Social Media; 2009; Sevilla, Spain p. 9-17.
- Vadileios L, Cristianini N. Tracking the flu pandemic by monitoring the social web. 2010 Presented at: 2nd IAPR Workshop on Cognitive Information Processing (CIP); June 14-16, 2010; Elba Island, Italy p. 411-416.
- Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *Am J Trop Med Hyg* 2012 Jan;86(1):39-45 [[FREE Full text](#)] [doi: [10.4269/ajtmh.2012.11-0597](https://doi.org/10.4269/ajtmh.2012.11-0597)] [Medline: [22232449](#)]
- Kim EK, Seok JH, Oh JS, Lee HW, Kim KH. Use of Hangeul Twitter to track and predict human influenza infection. *PLoS One* 2013;8(7):e69305 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0069305](https://doi.org/10.1371/journal.pone.0069305)] [Medline: [23894447](#)]
- Paul M, Dredze M. You are what you tweet: analyzing Twitter for public health. *Artificial Intelligence* 2011:265-272.
- Cheng Z, Caverlee J, Lee K. You are where you tweet: a content-based approach to geo-locating Twitter users. 2011 Presented at: 19th ACM International Conference on Information Knowledge Management; 2010; Scottsdale, AZ, USA p. 759-768.
- Achrekar H, Gandhe A, Lazarus R. Twitter improves seasonal influenza prediction. *HEALTHINF* 2012:61.
- Butler D. When Google got flu wrong. *Nature* 2013 Feb 14;494(7436):155-156. [doi: [10.1038/494155a](https://doi.org/10.1038/494155a)] [Medline: [23407515](#)]
- Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol* 2013 Oct;9(10):e1003256 [[FREE Full text](#)] [doi: [10.1371/journal.pcbi.1003256](https://doi.org/10.1371/journal.pcbi.1003256)] [Medline: [24146603](#)]
- Lazer D, Kennedy R, King G, Vespignani A. The parable of Google Flu: traps in big data analysis. *Science* 2014 Mar 13;343(6176):1203-1205. [doi: [10.1126/science.1248506](https://doi.org/10.1126/science.1248506)]
- Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One* 2010 Nov;5(11):e14118 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0014118](https://doi.org/10.1371/journal.pone.0014118)] [Medline: [21124761](#)]
- Culotta A. Detecting influenza epidemics by analyzing Twitter messages. 2010. URL: <http://arxiv.org/pdf/1007.4748.pdf> [accessed 2014-09-30] [[WebCite Cache ID 6S7zHUdmU](#)]
- Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. 2010 Presented at: KDD Workshop on Social Media Analytics; July 25, 2010; Washington, DC.

22. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One* 2011 May;6(5):e19467 [FREE Full text] [doi: [10.1371/journal.pone.0019467](https://doi.org/10.1371/journal.pone.0019467)] [Medline: [21573238](https://pubmed.ncbi.nlm.nih.gov/21573238/)]
23. Santos J, Sergio M. Predicting flu incidence from Portuguese tweets. In: IWBBIO Proceedings. 2013 Presented at: IWBBIO; March 18-20, 2013; Granada, Spain p. 11-18.
24. Broniatowski D, Paul M, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. *PLoS ONE* 2013;8(12):e83672. [doi: [10.1371/journal.pone.0083672](https://doi.org/10.1371/journal.pone.0083672)]
25. MappyHealth - Social Health Insights. URL: <http://nowtrending.hhs.gov/> [accessed 2014-03-21] [WebCite Cache ID [6OFZ65y1B](https://www.webcitation.org/6OFZ65y1B)]
26. Fountin Germ Tracker Map. URL: <http://fount.in/m> [accessed 2014-03-21] [WebCite Cache ID [6OFZ95jmm](https://www.webcitation.org/6OFZ95jmm)]
27. Observatorio da Dengue. URL: <http://www.observatorio.inweb.org.br/dengue/conteudo/inicial> [accessed 2014-03-21] [WebCite Cache ID [6OFZF730E](https://www.webcitation.org/6OFZF730E)]
28. Infovigil. URL: <http://www.infodemiology.org/> [accessed 2014-09-28] [WebCite Cache ID [6Sw4BEnvu](https://www.webcitation.org/6Sw4BEnvu)]
29. SickWeather. URL: <http://www.sickweather.com/> [accessed 2014-03-21] [WebCite Cache ID [6OFZIVyCc](https://www.webcitation.org/6OFZIVyCc)]
30. New York City Department of Health and Mental Hygiene. URL: <http://www.nyc.gov/html/doh/flu/html/data/data.shtml> [accessed 2014-03-21] [WebCite Cache ID [6OFZSNGPx](https://www.webcitation.org/6OFZSNGPx)]
31. WebPlot Digitizer. URL: <http://sourceforge.net/projects/digitizer/files/Engauge%20Digitizer/digitizer-5.1/> [accessed 2014-03-21] [WebCite Cache ID [6OFZNLuZ2](https://www.webcitation.org/6OFZNLuZ2)]
32. Dredze M, Paul M, Bergsma S, Tran H. Carmen: a Twitter geolocation system with applications to public health. 2013 Presented at: AAAI Workshop on Expanding the Boundaries of Health Informatics Using AI (HIAD); 2013; Palo Alto, California.
33. Lamb A, Paul M, Dredze M. Separating fact from fear: tracking flu infections on Twitter. *North American Chapter of the Association for Computational Linguistics (NAACL) 2013*:789-795.
34. CrowdBreaks. URL: <http://www.crowdbreaks.com/> [accessed 2014-03-21] [WebCite Cache ID [6OFZWt8Dy](https://www.webcitation.org/6OFZWt8Dy)]
35. Kulldorff M, Heffernan R, Hartman J, Assunção RM, Mostashari F. A space-time permutation scan statistic for the early detection of disease outbreaks. *PLoS Medicine* 2005;2:216-224. [doi: [10.1371/journal.pmed.0020059](https://doi.org/10.1371/journal.pmed.0020059)]
36. Jung I, Kulldorff M, Klassen AC. A spatial scan statistic for ordinal data. *Stat Med* 2007 Mar 30;26(7):1594-1607. [doi: [10.1002/sim.2607](https://doi.org/10.1002/sim.2607)] [Medline: [16795130](https://pubmed.ncbi.nlm.nih.gov/16795130/)]
37. ArcGIS for Desktop, Release 10. Redlands, CA: Esri; 2011. URL: <http://www.esri.com/> [accessed 2014-09-30] [WebCite Cache ID [6Sz7mMXix](https://www.webcitation.org/6Sz7mMXix)]
38. Study results. North Brooklyn Epicenter URL: <https://www.google.com/maps/place/40%C2%B041'06.0%22N+73%C2%B058'58.8%22W/@40.6834315,-73.9849816,16z/data=!4m2!3m1!1s0x0:0x0> [accessed 2014-09-30] [WebCite Cache ID [6SzOFpnVa](https://www.webcitation.org/6SzOFpnVa)]
39. Gu H, Chen B, Zhu H, Jiang T, Wang X, Chen L, et al. Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *J Med Internet Res* 2014;16(1):e20 [FREE Full text] [doi: [10.2196/jmir.2911](https://doi.org/10.2196/jmir.2911)] [Medline: [24440770](https://pubmed.ncbi.nlm.nih.gov/24440770/)]
40. Bernardo T, Rajic A, Young I, Robiadek K, Pham M, Funk J. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *J Med Internet Res* 2013;15(7):e147 [FREE Full text] [doi: [10.2196/jmir.2740](https://doi.org/10.2196/jmir.2740)] [Medline: [23896182](https://pubmed.ncbi.nlm.nih.gov/23896182/)]
41. Nagel AC, Tsou MH, Spitzberg BH, An L, Gawron JM, Gupta DK, et al. The complex relationship of realspace events and messages in cyberspace: case study of influenza and pertussis using tweets. *J Med Internet Res* 2013;15(10):e237 [FREE Full text] [doi: [10.2196/jmir.2705](https://doi.org/10.2196/jmir.2705)] [Medline: [24158773](https://pubmed.ncbi.nlm.nih.gov/24158773/)]
42. Leetaru K, Wang S, Padmanabhan A, Shook E. Mapping the global Twitter heartbeat: the geography of Twitter. *FM* 2013 May 06;18(5):1-34 [FREE Full text] [doi: [10.5210/fm.v18i5.4366](https://doi.org/10.5210/fm.v18i5.4366)]
43. Burton SH, Tanner KW, Giraud-Carrier CG, West JH, Barnes MD. "Right time, right place" health communication on Twitter: value and accuracy of location information. *J Med Internet Res* 2012;14(6):e156 [FREE Full text] [doi: [10.2196/jmir.2121](https://doi.org/10.2196/jmir.2121)] [Medline: [23154246](https://pubmed.ncbi.nlm.nih.gov/23154246/)]
44. Sadilek A, Kautz H, Silenzio V. Predicting disease transmission from geo-tagged micro-blog data. 2012 Presented at: 26th AAAI Conference on Artificial Intelligence; 2012; Toronto.

Abbreviations

- A:** Awareness tweet
- AR:** auto-regressive
- GSQ:** Google Trends search query
- H:** High probability of tweet veracity
- I:** Infection tweet
- ILI-ED:** influenza-like illness emergency department visit
- L:** Low probability of tweet veracity

M: Medium probability of tweet veracity
MAPE: mean absolute percent error
NYC: New York City
O: Subject of tweet is “other” person
R: Relevant tweet
S: Subject of tweet is self

Edited by G Eysenbach; submitted 22.03.14; peer-reviewed by M Kang, MH Tsou, O Leal Neto; comments to author 15.06.14; revised version received 08.08.14; accepted 30.08.14; published 20.10.14.

Please cite as:

*Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, Brownstein JS
A Case Study of the New York City 2012-2013 Influenza Season With Daily Geocoded Twitter Data From Temporal and Spatiotemporal Perspectives
J Med Internet Res 2014;16(10):e236
URL: <http://www.jmir.org/2014/10/e236/>
doi: [10.2196/jmir.3416](https://doi.org/10.2196/jmir.3416)
PMID: [25331122](https://pubmed.ncbi.nlm.nih.gov/25331122/)*

©Ruchit Nagar, Qingyu Yuan, Clark C Freifeld, Mauricio Santillana, Aaron Nojima, Rumi Chunara, John S Brownstein. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Evaluating the Process of Online Health Information Searching: A Qualitative Approach to Exploring Consumer Perspectives

Alexander S Fiksdal¹, MA; Ashok Kumbamu¹, PhD; Ashutosh S Jadhav²; Cristian Cocos³, PhD; Laurie A Nelsen³; Jyotishman Pathak², PhD; Jennifer B McCormick^{1,4,5}, MPP, PhD

¹Mayo Clinic, Biomedical Ethics Program, Rochester, MN, United States

²Mayo Clinic, Biomedical Statistics and Informatics, Rochester, MN, United States

³Mayo Clinic, Global Business Solutions, Rochester, MN, United States

⁴Mayo Clinic, General Internal Medicine, Rochester, MN, United States

⁵Mayo Clinic, Health Care Policy and Research, Rochester, MN, United States

Corresponding Author:

Jennifer B McCormick, MPP, PhD

Mayo Clinic

General Internal Medicine

200 First Street SW

Rochester, MN,

United States

Phone: 1 507 293 0185

Fax: 1 507 538 0850

Email: McCormick.JB@mayo.edu

Abstract

Background: The Internet is a common resource that patients and consumers use to access health-related information. Multiple practical, cultural, and socioeconomic factors influence why, when, and how people utilize this tool. Improving the delivery of health-related information necessitates a thorough understanding of users' searching-related needs, preferences, and experiences. Although a wide body of quantitative research examining search behavior exists, qualitative approaches have been under-utilized and provide unique perspectives that may prove useful in improving the delivery of health information over the Internet.

Objective: We conducted this study to gain a deeper understanding of online health-searching behavior in order to inform future developments of personalizing information searching and content delivery.

Methods: We completed three focus groups with adult residents of Olmsted County, Minnesota, which explored perceptions of online health information searching. Participants were recruited through flyers and classified advertisements posted throughout the community. We audio-recorded and transcribed all focus groups, and analyzed data using standard qualitative methods.

Results: Almost all participants reported using the Internet to gather health information. They described a common experience of searching, filtering, and comparing results in order to obtain information relevant to their intended search target. Information saturation and fatigue were cited as main reasons for terminating searching. This information was often used as a resource to enhance their interactions with health care providers.

Conclusions: Many participants viewed the Internet as a valuable tool for finding health information in order to support their existing health care resources. Although the Internet is a preferred source of health information, challenges persist in streamlining the search process. Content providers should continue to develop new strategies and technologies aimed at accommodating diverse populations, vocabularies, and health information needs.

(*J Med Internet Res* 2014;16(10):e224) doi:[10.2196/jmir.3341](https://doi.org/10.2196/jmir.3341)

KEYWORDS

Internet; information seeking behavior; consumer health information; qualitative research

Introduction

In recent years, the quantity and quality of health information available on the Internet has increased substantially. As access to reliable, affordable, high-speed Internet access increases, the percentage of people using the Internet to search and subsequently learn from health-related information continues to grow rapidly as well. In the current climate of rising costs of health care in the United States, the role of freely available health care information is becoming more central to patients, their families and friends, and even health care providers. In order to improve the delivery of content, researchers and scientists must first develop a thorough understanding of the searching-related needs and experiences of users.

Recent studies have shed light on why and how consumers search for health information on the Internet [1-3]. In a recent 2013 survey conducted by the Pew Internet Project, 72% of respondents reported using the Internet to look for health information within the past year, with the most commonly researched topics being focused on specific diseases or conditions, treatments or procedures, and searching for doctors or other health professionals [1]. Although many people (35% of those surveyed by Pew) use the Internet to learn more about a specific symptom or medical condition they or someone else might have, clinicians and/or family and friends remain a central resource when help is needed regarding a serious health issue [2,4]. The elderly in particular are more likely to trust “living sources” of information, rather than the Internet [3]. Even among Internet users, health information is often understood in a social context. For example, 26% of Internet users reported watching or reading content related to someone else’s personal experience with a medical or health-related issue within the last 12 months [1].

Health information seeking behavior depends on a variety of factors including subjective factors (eg, intent for the search, experience in using and searching the Internet, and information preferences [5,6]) and socioeconomic factors (eg, age group, income level, education level, etc [4,7,8]). Research shows that women are more likely than men to search for health information [9] and online health consumers tend to be more educated, earn more, and have high-speed Internet access at home and at work [5,10,11]. Although low-income individuals do use the Internet, some may have difficulty distinguishing between low and high quality information [12]. Additionally, low-income disabled and homebound adults show lower rates of Internet use overall [13]. Further, our preliminary results from another study indicate that online health information seeking behavior differs significantly compared to general information searching. In particular, our data suggests that health-related queries are typically longer (ie, more words) and contextual in nature compared to general queries [14]. Also, health-related queries have higher rates of misspelled words that are typically corrected by “auto-completion” features available universally in all Web search engines such as Google and Bing [14].

There are various motivating factors for health information searching on the Internet. Aside from trying to learn more about a symptom or disorder specifically relevant to the person

searching, half of online health information research is on behalf of a friend or relative [15]. Additionally, searching is often used to track specific health-related factors. For example, 60% of adults reported tracking their weight diet or exercise routine online, and 33% reported tracking specific health indicators or symptoms such as blood pressure, blood sugar, headaches, or sleep patterns [1].

A large proportion of the population uses the Internet to search for health information, and their motivations for doing so are varied [1-3]. This complex situation, along with an educationally and culturally heterogeneous population, has resulted in a barrier in the process of gathering and interpreting health information. In this context, the preferred vocabulary within and between different groups of people can differ significantly, often resulting in a variety of words being used to describe the same concept or medical condition [16-18]. Knowledge gaps can then emerge between patients and providers. One possible strategy for addressing such gaps involves developing consumer-focused vocabularies and associated infrastructure for health information retrieval that can act as an interface between parties [19]. Before such vocabularies and technologies can be developed, researchers and scientists must have a thorough understanding of the current state of online health information searching. While a large body of survey-based research has been conducted regarding this subject [1-3,20], qualitative research provides a unique perspective that can play a valuable role in informing future research and technological developments. The aim of the current study was to engage in in-depth discussions with community members about their health-related searching activities. All the study participants are residents of Olmsted County, Minnesota (MN), and are either Mayo Clinic patients, employees, or at least have one family member at home who is a patient or employee.

Methods

Study Participants and Recruitment

To better understand health information searching behavior and its implications for health and well-being of community members, we conducted three 90-minute focus groups of 5 to 6 individuals over the course of a 2-month span. We targeted adult, English-speaking members of the Olmsted County, MN community (where Mayo Clinic is located) and Mayo Clinic patient, employees, and family visitors. We recruited participants using flyers and online classifieds ads distributed throughout the Rochester, MN community and within Mayo Clinic. [Table 1](#) summarizes basic characteristics of participants. Participants were provided a modest financial remuneration for participating in the study.

Moderators (JM and AK) trained in qualitative methodology facilitated discussions about the attitudes and experiences of participants related to searching for health information on the Internet. Moderators used a semi-structured moderator guide to facilitate discussion and the guide covered four major aspects: (1) participants’ perception and understanding of health care information, (2) the process of information collection on the Internet, (3) understanding and usage of information, and (4) implications of health care information for their health and

well-being. Participants were asked about their thoughts and the connotations surrounding each of these themes. Oral consent was obtained from all participants. This study was approved by the Institutional Review Board at Mayo Clinic (IRB #12-005476).

Prior to participating in the focus groups, participants completed an anonymous questionnaire that included questions assessing basic demographic information and previously used sources of health information. All focus groups were audio-recorded, transcribed, and de-identified.

Table 1. Characteristics of patients (n=19).

Characteristic	n (%)
Age, mean (SD; range)	43.26 (17.0; 22-73)
Sex	
Male	5 (26%)
Female	14 (74%)
Race	
White	15 (79%)
Black or African American	0 (0%)
Asian	4 (21%)
Highest level of education	
High school or GED	0 (0%)
Community or Jr. College	3 (16%)
Four-year college	3 (16%)
Graduate school	13 (68%)
Yearly household income (US\$)	
Less than \$15,000	0 (0%)
\$15,000-\$35,000	2 (11%)
\$35,001-\$55,000	9 (47%)
\$55,001-\$75,000	4 (21%)
\$75,001-\$100,000	0 (0%)
Over \$100,000	1 (5%)
Prefer not to answer	3 (16%)
Prior sources used to get health information	
Health care providers	19 (100%)
Family/friends	15 (79%)
Organizations/support groups	6 (32%)
Internet	18 (95%)
Books/pamphlets	15 (79%)
Other	1 (5%)
Prior participation in research	
Yes	4 (21%)
No	15 (79%)

Data Collection and Analysis

All team members read de-identified transcripts and developed a codebook through an iterative process [21]. Using the codebook, two members of the team independently coded the transcripts in NVivo, a qualitative software application. The data were then analyzed using a grounded-theory approach (NVivo qualitative data analysis software; QSR International

Pty Ltd. Version 10, 2012). Coding inconsistencies were discussed and resolved through consensus, with the input of a third team member when necessary.

Results

Overview

Participants candidly discussed how they used the Internet to search for health information. Through these discussions, several themes related to health motivations, content preferences, and practical applications of searching emerged. Below we summarize these data in the context of three major themes: motivations for searching, searching strategies and techniques, and information content preferences.

Motivations for Online Health Searching

A variety of factors play a role in initiating online searches for health information. The motivations that our participants described generally fell into three main areas: (1) symptom troubleshooting, (2) searching to enhance a clinic visit, and (3) proxy searching.

Perhaps the most common motivation for everyday searching is a phenomenon that could be called “symptom troubleshooting”. With commercial online resources and other government or hospital/university-based sites that provide free, anonymous, and immediate information, many individuals’ first stop to learn more about a specific symptom is the Internet. A participant from Focus Group (FG) #3 mentioned: “For me, it was very important when I think I have a symptom, the first place I look is the Internet, especially to search for the symptoms”.

Once a particular symptom or disorder of interest is identified, participants reported that the Internet made it very easy to get more detailed information to help identify underlying causes. As a participant of FG #3 explained: “For instance if I have a pain in my foot, I am going to start looking for...information that might specify if it’s in the heel or in the toe...then I search [for] why [I have] the symptom or, if I know what I have, then I might search...to see if I can match the symptoms to that”.

Using the Internet provided a quick and easy way to troubleshoot symptoms; however, there are certain situations where using the Internet is more likely. One participant explained that the Internet is especially more convenient for superficial symptoms: “You can’t just go find a doctor somewhere and be ‘hey, can you look at this rash on my leg’ because I hear doctors hate that” [FG #1]. The Internet provides a level of anonymity that may be helpful in situations where individuals perceive their problems to be bothersome or nuisances to doctors.

Participants often cited practical reasons related to time and money when describing their motivations for turning to the Internet for medical information or advice. One participant explained that although consulting a professional in person can be preferable, “especially when you are very concerned about your symptoms”, in other cases, as he stated, “at 9:00 at night you are not going to be able to call the doctor” [FG #3]. Another participant in FG #1 also echoed a similar sentiment: “It can’t be readily available, you may have to make a doctor’s appointment and that could take a while...and cost money and financially that might hold you back too; something that a fast care isn’t going to be able to fix”.

For non-serious medical issues, participants were generally comfortable using the Internet as a troubleshooting tool. Once a health care provider is involved, however, searching assumes a different role. In this context, participants reported using Internet searching as a means to enhance a clinic visit and be more well-prepared and well-informed during the entire health care experience with their providers. In these situations, Internet searching proved to be a valuable tool in preparing for the clinic visit. As one participant in FG #1 explained, Internet searching allowed her to walk into a surgery consultation armed with a prior understanding of possible procedures: “I specifically knew all the three main surgeries; I knew what I liked from them, what I didn’t like of them”.

This online preparation gave her the information and ability to “say what about this, what about that, why are we doing this, why are we doing that?” [FG #1]. Participants agreed that such preparation facilitates “a more enriched experience” [FG #1] and allows patients to “become more knowledgeable” and “ask better questions” to providers [FG #2]. This participant goes on to explain how such a dynamic increases communication and education and “builds the patient/provider relationship”; “If you are taking an interest in what it is you have and asking the kind of questions that allow them to further educate you, I think that shows a real interest” [FG #2].

Another participant expanded on this idea and explained how an enriched patient/provider relationship involves more than developing a healthy rapport and can actually improve health outcomes in certain situations: “I mean my mom had a weird thyroid thing and she was all over the Internet, and still is, but she would bring stuff to her doctor and she actually like did solve some mysterious things and she gave stuff to her doctor and her doctor I think is a great doctor but there is so much information and the doctors don’t get it all” [FG #2].

In the previous example, the participant’s mother used the Internet for two of the main motivations that emerged from our focus groups: to troubleshoot a thyroid condition and to enhance her visits with her doctors. Although this participant’s mother was able to do the searching and advocating on her own, many participants had parents, grandparents, or other family members who were not as comfortable or capable. These situations highlight the third main motivation for searching that our participants discussed: searching for someone else, or proxy searching. All of the focus groups had participants who reported searching on behalf of someone else. For many, it was a frequent occurrence.

Computer literacy was often cited as a main reason for proxy searching, as many participants had relatives who were “afraid of using it [computers and the Internet]” [FG #1]. However, proxy searching was also a useful tactic when the individual searching sought to protect their relative from additional emotional burdens, even when the relative was computer literate. One focus group participant explained: “Well, I have done searches for my parents before... When I looked up stuff [about] breast cancer on the Internet, [I told them] do not look it up because you’re going to be scared. As a third person, even though she is my mom, I know how to decide and to remove

myself from the situation, but she is not going to be able to do that" [FG #3].

Searching Strategies and Techniques

In terms of the actual mechanics of searching, participants described using a common set of steps and procedures that began with commonly used search engines, continued to shop around for information from various sources, and ended with information saturation and exhaustion.

Regardless of the underlying motivations for searching, almost all searches shared a common starting point from an online Web search engine: Google. Ease of use—"you can ask the most stupidest questions and have a pretty good shot of getting an answer" [FG #1]—and quality of results—"[Google] brings up the most variety of answers" [FG #1]—were the primary reasons for choosing Google cited by our participants.

Although Google is by far the most common first step to searching, its main use is simply as a tool to reach other sites. One participant mentioned: "Google's just a way to get there" [FG #1]. Another participant expanded on this view, adding "I agree. I am not putting my trust in Google; I am only putting my trust that Google is going to give me a variety. My trust is actually embedded only in the searches I click, it is just the outlet to get me there, it is just the bridge" [FG #1].

Once Google supplied a list of relevant sites to visit, most participants reported visiting many sites in order to satisfy their searching demands. This technique allows participants to "shop around and have multiple sources" without having to use exact phrasing [FG #1]. The information shopping process described by participants often included multiple side-by-side comparisons. One participant mentioned: "Because you can have multiple open window task bars and tabs on the Web browser, I open every single one on the first page in each of the task bars and compare all of them" [FG #2].

This technique facilitated the information shopping experience and gives greater confidence in results because "you get as much information as you can if [all the websites] have the same information" [FG #2]. Many participants used the tabs function of Web browsers to compare multiple websites at once.

Participants described a common sequence of events that led to the termination of the search process. As the comparing and filtering process of multiple websites progresses, participants reported that eventually "all the information is basically the same" [FG #1]. Although another participant acknowledged that "there are always additional links to go to" [FG #1], other participants explained that once results became irrelevant to their original search query it was time to stop the search process. One participant explained: "If you go down to the 17th, 20th, 30th option under Google, you find that what you are looking for is the 30th degree of separation. It is just not as relevant to what it is you are trying to research anymore" [FG #1].

Some participants also reported a sense of being "lost" or "completely forgetting where you started", especially in cases of performing broad searches. The resulting confusion can lead to becoming "unmotivated" to continue searching, even if the original query has not been resolved [FG #1].

In addition to information saturation, subjective fatigue was an indicator participants described as a reason for ending the search process. After a long, drawn-out search process, participants reported getting "tired with the screens" and feeling "exhausted" [FG #1]. Another participant compared the process to shopping: "If you know what you want, you can go to ten different places to try to find that one thing, but after a while...you are going to be hitting your head against the wall...it gets exhausting" [FG #1].

Ultimately, the participants described searching for health-related information as a rigorous process of comparing and contrasting various sources against personalized criteria based on need and individual appraisal of reputation. This filtering process generally continues until the results become repetitive and/or the searcher becomes fatigued.

Content Preferences

Major search engines can easily produce thousands of results for any given query. How then do patients and consumers select which websites to gather health-related information? Although every search is unique, participants overwhelmingly preferred sites based on two main factors: reputation and advertising (or lack thereof).

Participants often commented that they "tend to go for the sites that are most reputable" [FG #1]. While the importance of reputation applied to all websites, regardless if they were related to health, participants also reported placing a higher standard of quality on health-related information. As one participant explained, "Health is unlike any other consumer type of website...I take it to a totally different level. I want to have the best, you only have one body" [FG #1]. Making sure they had "the best" gave participants comfort in knowing they were receiving accurate information. Often "the best" is synonymous with dealing with a "reputable institution", which is in turn largely influenced by branding. One participant explained: "When you are dealing with a company, an organization that has a good reputation, then you feel more confident that you are getting the right information" [FG #3].

In addition to pure name recognition, participants reported that institutions "earn trust...through publications, research, and education" [FG #2]. Additionally, "how [websites or institutions] are ranked" or if they are "well known" contributed to participants' conception of reputation [FG #2]. Finally, participants were more likely to view sources of health information as reputable if they were domestic. As one participant explained, "I would rely more heavily on those [domestic] institutions than a foreign hospital that may be quite good but is somewhere outside of the United States" [FG #2].

While reputation played a major role in determining which websites to trust for our participants, advertising and commercial interests often dissuaded them. Almost all of our participants reported avoiding websites that had visible advertising or were obviously profit-oriented. As one participant explained, "If I see ads, I question the motivation for providing information that they have" [FG #1]. Another participant explained the aversion in the context of a wider trend of commercialization of medicine: "I think for me it scares me how, and I suppose

this could go onto a variety of different things, but it scares me how medicine has transformed into such a consumer-driven place” [FG #1].

Most of our participants shared distaste for commercial interests in their searching behavior; however, in some cases it had more to do with the perception of profit-driven motivations rather than the true nature of the business or organization. In response to a question regarding whether or not participants thought that MayoClinic.com, the commercial consumer health information portal owned and maintained by Mayo Clinic, was a “commercial” website, one participant responded, “Well, you don’t see a lot of advertising on the Mayo site...I don’t see a lot going on the sides all the way down the page flashing at me, I don’t have a lot of popups that come at me” [FG #1].

Although Mayo Clinic does indeed utilize advertising on the website, the combined name recognition, familiarity, and subtle nature of advertisements was enough to retain credibility for many of our participants. We acknowledge that there might be an inherent bias in this finding since the study participants were either Mayo Clinic patients, employees, or at least have one family member at home who is a patient or employee.

Discussion

Principal Findings

Our goal in collecting these qualitative data was to better understand how consumers use and search for health information on the Internet to inform the development of more personalized health information searching and delivery applications. The participants in this study described a common experience of searching for health information that largely mirrors recent large-scale survey data. Most of our participants see the Internet as a potentially valuable tool to find information about health and medical conditions; yet, they did point to the challenge of efficiently addressing their particular needs given the vast amounts of information. This reflects the challenge of streamlining and personalizing information for a user base that is diverse both in terms of individual background and need. The data presented here, particularly in the context of content preferences and searching techniques, may be beneficial to researchers and content providers as they develop new strategies for delivering health information.

Many participants shared examples of how they use information they found through Internet searches in their efforts to enhance their interactions with their health care providers. Examining these data in the context of increasing health costs and physician time constraints provides valuable insight into the challenges and opportunities consumers and physicians will encounter in years ahead. Many of our participants reported using Internet health searching as a means of enhancing clinic visits, either through preparation or post-appointment follow-up. Some concerns exist regarding how doctors may react to patients introducing health information gathered from the Internet into the exam room, and indeed previous research has indicated that some physicians view such occurrences negatively [22,23]. Patients, on the other hand, tend to view Internet health searching as an additional resource to complement the still

highly valued patient/physician relationship [24,25]. Our data also support this view of the patient perspective, as our participants viewed online health searching as a means to “build the doctor-patient relationship” [FG #2]. How physicians respond likely depends on physician communication skills and whether or not the physician feels challenged [26]. The participant experiences and opinions described here are largely from a patient perspective and are largely positive in the context of using health information from the Internet to enhance visits. These perspectives may be useful in framing future research focused on physician perspectives on using such information in office visits.

Recently, the amount of time doctors spend in front of patients has received attention in the media [27,28]. Having patients armed with information and questions prior to office visits may help improve care in the current realities of decreased face time with doctors, which today can be as low as 8 minutes on average [28]. This of course necessitates that the information patients gather be of high quality. Indeed, research suggests the quality of information that patients present ultimately determines its effect on the patient/physician relationship; while accurate information can be helpful, inaccurate information may be harmful [29]. Our future work will therefore focus on ways to develop consumer health information technology solutions to facilitate the transmission of accurate, trustworthy, validated information to consumers to ensure that online health information searching enhances, rather than hinders, care.

Limitations

This study contained a few important limitations. Due to recruitment constraints, our study population was limited to adults within Olmsted County, MN. All participants were either employees or were family members of employees and patients at Mayo Clinic, where the study took place. Additionally, our sample was highly educated, with all participants having attained at least a community college degree, and 68% having completed graduate school. We were therefore unable to explore the perspectives of a more diverse population. It is also important to consider our choice of study design when interpreting the data we presented. In this study, we used qualitative approaches such as grounded theory and focus groups method for data collection and analysis. These qualitative methods allow us to contextualize participants’ understandings and experiences, to track variations in how concepts are understood, and to uncover novel findings that may warrant further investigation [30]. In this way, we are able to make, as Giacomini and Cook describe, an “empirically-based contribution to ongoing dialogue” [31]. The overarching goal of qualitative research is to explore and describe *particularities* of a social phenomenon rather than producing generalizable results. But, findings from a small sample size in a qualitative research can help developing hypothesis for a quantitative study to produce generalizable findings from a larger sample size. Our study participants were recruited from a limited subset of individuals that was readily accessible in a community dominated by the health care industry. In doing so, our goal is not to present data that can or should be generalized to a wider population, but rather to explore pertinent issues with a level of depth that is not possible with standard quantitative (and generalizable) methodologies. Indeed,

we cannot claim that the experiences described here are representative of all Internet users; however, they can inform the development of future work and research in areas of streamlining content delivery and patient/physician interaction.

Conclusions

We conducted this qualitative study to gain a deeper understanding of search behavior in order to inform future technological developments in personalizing online information searching and content delivery. Although the Internet was a preferred source of health information for almost all of our

participants, from a consumer and patient perspective challenges persist in streamlining the process of identifying reliable and high quality content that also matches the intended search target of the user. Our participants described a current search paradigm consisting of drawn-out user-driven comparisons of content obtained from multiple sources of varying quality and unverified validity. As consumers continue to use information gathered from the Internet to enhance their interactions with health care providers, new strategies for delivering health information on the Internet must be developed that accommodate diverse backgrounds and clinical needs.

Acknowledgments

The authors would like to acknowledge Thomas Suther for providing access to the Mayo Clinic digital usability laboratory where one of the focus groups was conducted. This project was supported by the Mayo Clinic and by Grant Number UL1 TR000135 from the National Center for Advancing Translational Sciences (NCATS).

Conflicts of Interest

None declared.

References

1. Fox S. Pew Internet & American Life Project report. 2013. Pew Internet: Health URL: <http://www.pewinternet.org/fact-sheets/health-fact-sheet/> [accessed 2014-02-21] [WebCite Cache ID 6NYqjGvWn]
2. Gutierrez N, Kindratt TB, Pagels P, Foster B, Gimpel NE. Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. *J Community Health* 2014 Feb;39(1):83-89. [doi: [10.1007/s10900-013-9742-5](https://doi.org/10.1007/s10900-013-9742-5)] [Medline: [23900880](https://pubmed.ncbi.nlm.nih.gov/23900880/)]
3. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. *Comput Inform Nurs* 2013 Nov;31(11):547-553 [FREE Full text] [doi: [10.1097/01.NCN.0000432131.92020.42](https://doi.org/10.1097/01.NCN.0000432131.92020.42)] [Medline: [23974574](https://pubmed.ncbi.nlm.nih.gov/23974574/)]
4. Cutilli CC. Seeking health information: what sources do your patients use? *Orthop Nurs* 2010;29(3):214-219. [doi: [10.1097/NOR.0b013e3181db5471](https://doi.org/10.1097/NOR.0b013e3181db5471)] [Medline: [20505493](https://pubmed.ncbi.nlm.nih.gov/20505493/)]
5. Higgins OSJ, Barry MM, Dmegan C. A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective. *European Centre for Disease Control* 2011 Oct:1-16 [FREE Full text]
6. Lorence DP, Park H, Fox S. Assessing health consumerism on the Web: a demographic profile of information-seeking behaviors. *J Med Syst* 2006 Aug;30(4):251-258. [Medline: [16978004](https://pubmed.ncbi.nlm.nih.gov/16978004/)]
7. Drentea P, Goldner M, Cotten S, Hale T. The association among gender, computer use and online health searching, and mental health. *Information, Communication & Society* 2008 Jun;11(4):509-525. [doi: [10.1080/13691180801999019](https://doi.org/10.1080/13691180801999019)]
8. 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](https://doi.org/10.1111/j.1525-1497.2006.00622.x)] [Medline: [16995892](https://pubmed.ncbi.nlm.nih.gov/16995892/)]
9. 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](https://doi.org/10.2196/jmir.1035)] [Medline: [19275980](https://pubmed.ncbi.nlm.nih.gov/19275980/)]
10. Wangberg SC, Andreassen HK, Prokosch HU, Santana SM, Sorensen T, Chronaki CE. Relations between Internet use, socio-economic status (SES), social support and subjective health. *Health Promot Int* 2008;23(1):70-77. [doi: [10.1093/heapro/dam039published](https://doi.org/10.1093/heapro/dam039published)]
11. 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](https://doi.org/10.2196/jmir.1023)] [Medline: [19017584](https://pubmed.ncbi.nlm.nih.gov/19017584/)]
12. Knapp C, Madden V, Wang H, Sloyer P, Shenkman E. Internet use and eHealth literacy of low-income parents whose children have special health care needs. *J Med Internet Res* 2011;13(3):e75 [FREE Full text] [doi: [10.2196/jmir.1697](https://doi.org/10.2196/jmir.1697)] [Medline: [21960017](https://pubmed.ncbi.nlm.nih.gov/21960017/)]
13. 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](https://doi.org/10.2196/jmir.2645)] [Medline: [23639979](https://pubmed.ncbi.nlm.nih.gov/23639979/)]
14. Jadhav A, Andrews D, Fiksdal A, Kumbamu A, McCormick JB, Misitano A, et al. Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal. *J Med Internet Res* 2014;16(7):e160 [FREE Full text] [doi: [10.2196/jmir.3186](https://doi.org/10.2196/jmir.3186)] [Medline: [25000537](https://pubmed.ncbi.nlm.nih.gov/25000537/)]

15. Sadasivam RS, Kinney RL, Lemon SC, Shimada SL, Allison JJ, Houston TK. Internet health information seeking is a team sport: analysis of the Pew Internet Survey. *Int J Med Inform* 2013 Mar;82(3):193-200. [doi: [10.1016/j.ijmedinf.2012.09.008](https://doi.org/10.1016/j.ijmedinf.2012.09.008)] [Medline: [23149121](https://pubmed.ncbi.nlm.nih.gov/23149121/)]
16. Smith CA. Nursery, gutter, or anatomy class? Obscene expression in consumer health. 2009 Presented at: AMIA Annual Symposium; 2007; Chicago, Illinois p. 676-680.
17. Keselman A, Smith CA, Divita G, Kim H, Browne AC, Leroy G, et al. Consumer health concepts that do not map to the UMLS: where do they fit? *J Am Med Inform Assoc* 2008;15(4):496-505 [FREE Full text] [doi: [10.1197/jamia.M2599](https://doi.org/10.1197/jamia.M2599)] [Medline: [18436906](https://pubmed.ncbi.nlm.nih.gov/18436906/)]
18. Zielstorff RD. Controlled vocabularies for consumer health. *J Biomed Inform* 2003;36(4-5):326-333 [FREE Full text] [Medline: [14643728](https://pubmed.ncbi.nlm.nih.gov/14643728/)]
19. Seedorff M, Peterson KJ, Nelsen LA, Cocos C, McCormick JB, Chute CG, et al. Incorporating expert terminology and disease risk factors into consumer health vocabularies. 2013 Presented at: Pacific Symposium on Biocomputing; 2013; Kohala Coast, Hawaii p. 421-432 URL: http://psb.stanford.edu/psb-online/proceedings/psb13/abstracts/2013_p421.html
20. Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. *J Health Commun* 2010;15 Suppl 3:279-293. [doi: [10.1080/10810730.2010.522700](https://doi.org/10.1080/10810730.2010.522700)] [Medline: [21154099](https://pubmed.ncbi.nlm.nih.gov/21154099/)]
21. Corbin JM, Strauss AL. Basics of qualitative research: techniques and procedures for developing grounded theory. Los Angeles, California: Sage Publications, Inc; 2008:1-379.
22. Hamann J, Mendel R, Bühner M, Kissling W, Cohen R, Knipfer E, et al. How should patients behave to facilitate shared decision making--the doctors' view. *Health Expect* 2012 Dec;15(4):360-366. [doi: [10.1111/j.1369-7625.2011.00682.x](https://doi.org/10.1111/j.1369-7625.2011.00682.x)] [Medline: [21624024](https://pubmed.ncbi.nlm.nih.gov/21624024/)]
23. Ahmad F, Hudak PL, Bercovitz K, Hollenberg E, Levinson W. Are physicians ready for patients with Internet-based health information? *J Med Internet Res* 2006;8(3):e22 [FREE Full text] [doi: [10.2196/jmir.8.3.e22](https://doi.org/10.2196/jmir.8.3.e22)] [Medline: [17032638](https://pubmed.ncbi.nlm.nih.gov/17032638/)]
24. Kivits J. Informed patients and the internet: a mediated context for consultations with health professionals. *J Health Psychol* 2006 Mar;11(2):269-282. [doi: [10.1177/1359105306061186](https://doi.org/10.1177/1359105306061186)] [Medline: [16464924](https://pubmed.ncbi.nlm.nih.gov/16464924/)]
25. Stevenson FA, Kerr C, Murray E, Nazareth I. Information from the Internet and the doctor-patient relationship: the patient perspective--a qualitative study. *BMC Fam Pract* 2007;8:47 [FREE Full text] [doi: [10.1186/1471-2296-8-47](https://doi.org/10.1186/1471-2296-8-47)] [Medline: [17705836](https://pubmed.ncbi.nlm.nih.gov/17705836/)]
26. Murray E, Lo B, Pollack L. The impact of health information on the internet on the physician-patient relationship: patient perceptions. *Arch Intern Med* 2003;163(14):1727-1734. [doi: [10.1001/archinte.163.14.1727published](https://doi.org/10.1001/archinte.163.14.1727published)] [Medline: [12885689](https://pubmed.ncbi.nlm.nih.gov/12885689/)]
27. Chen PW. For New Doctors, 8 Minutes Per Patient. 2013. URL: http://well.blogs.nytimes.com/2013/05/30/for-new-doctors-8-minutes-per-patient/?_php=true&_type=blogs&_r=1 [accessed 2014-09-15] [WebCite Cache ID [6ScJM3VeK](https://www.webcitation.org/6ScJM3VeK)]
28. Block L, Habicht R, Wu AW, Desai SV, Wang K, Silva KN, et al. In the wake of the 2003 and 2011 duty hours regulations, how do internal medicine interns spend their time? *J Gen Intern Med* 2013 Aug;28(8):1042-1047 [FREE Full text] [doi: [10.1007/s11606-013-2376-6](https://doi.org/10.1007/s11606-013-2376-6)] [Medline: [23595927](https://pubmed.ncbi.nlm.nih.gov/23595927/)]
29. Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the Internet on health care and the physician-patient relationship: national U.S. survey among 1,050 U.S. physicians. *J Med Internet Res* 2003;5(3):e17 [FREE Full text] [doi: [10.2196/jmir.5.3.e17](https://doi.org/10.2196/jmir.5.3.e17)] [Medline: [14517108](https://pubmed.ncbi.nlm.nih.gov/14517108/)]
30. Hull S, Taylor H, Kass N. Qualitative Methods. In: Sugarman J, Sulmasy DP, editors. *Methods in Medical Ethics*. Washington, DC: Georgetown University Press; 2001:146-168.
31. Giacomini M, Cook DJ. Users' guides to the medical literature: XXIII. Qualitative research in health care B. What are the results and how do they help me care for my patients? Evidence-Based Medicine Working Group. *JAMA* 2000 Jul 26;284(4):478-482. [Medline: [10904512](https://pubmed.ncbi.nlm.nih.gov/10904512/)]

Edited by G Eysenbach; submitted 17.03.14; peer-reviewed by M Conway, S Koch-Weser; comments to author 17.04.14; revised version received 17.06.14; accepted 10.07.14; published 07.10.14.

Please cite as:

Fiksdal AS, Kumbamu A, Jadhav AS, Cocos C, Nelsen LA, Pathak J, McCormick JB

Evaluating the Process of Online Health Information Searching: A Qualitative Approach to Exploring Consumer Perspectives

J Med Internet Res 2014;16(10):e224

URL: <http://www.jmir.org/2014/10/e224/>

doi: [10.2196/jmir.3341](https://doi.org/10.2196/jmir.3341)

PMID: [25348028](https://pubmed.ncbi.nlm.nih.gov/25348028/)

©Alexander S Fiksdal, Ashok Kumbamu, Ashutosh S Jadhav, Cristian Cocos, Laurie A Nelsen, Jyotishman Pathak, Jennifer B McCormick. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Tablet, Web-Based, or Paper Questionnaires for Measuring Anxiety in Patients Suspected of Breast Cancer: Patients' Preferences and Quality of Collected Data

Maarten W Barentsz¹, MD; Hester Wessels¹, PhD; Paul J van Diest¹, MD, PhD; Ruud M Pijnappel¹, MD, PhD; Cees Haaring¹; Carmen C van der Pol¹, MD; Arjen J Witkamp¹, MD, PhD; Maurice A van den Bosch¹, MD, PhD; Helena M Verkooijen¹, MD, PhD

University Medical Center Utrecht, Utrecht, Netherlands

Corresponding Author:

Maarten W Barentsz, MD

University Medical Center Utrecht

Heidelberglaan 100

Utrecht, 3584 CX

Netherlands

Phone: 31 887556689

Fax: 31 302581098

Email: m.barentsz@umcutrecht.nl

Abstract

Background: Electronic applications are increasingly being used in hospitals for numerous purposes.

Objective: Our aim was to assess differences in the characteristics of patients who choose paper versus electronic questionnaires and to evaluate the data quality of both approaches.

Methods: Between October 2012 and June 2013, 136 patients participated in a study on diagnosis-induced stress and anxiety. Patients were asked to fill out questionnaires at six different moments during the diagnostic phase. They were given the opportunity to fill out the questionnaires on paper or electronically (a combination of tablet and Web-based questionnaires). Demographic characteristics and completeness of returned data were compared between groups.

Results: Nearly two-thirds of patients (88/136, 64.7%) chose to fill out the questionnaires on paper, and just over a third (48/136, 35.3%) preferred the electronic option. Patients choosing electronic questionnaires were significantly younger (mean 47.3 years vs mean 53.5 in the paper group, $P=.01$) and higher educated ($P=.004$). There was significantly more missing information (ie, at least one question not answered) in the paper group during the diagnostic day compared to the electronic group (using a tablet) (28/88 vs 1/48, $P<.001$). However, in the week after the diagnostic day, missing information was significantly higher in the electronic group (Web-based questionnaires) compared to the paper group (41/48 vs 38/88, $P<.001$).

Conclusions: Younger patients and patients with a higher level of education have a preference towards filling out questionnaires electronically. In the hospital, a tablet is an excellent medium for patients to fill out questionnaires with very little missing information. However, for filling out questionnaires at home, paper questionnaires resulted in a better response than Web-based questionnaires.

(*J Med Internet Res* 2014;16(10):e239) doi:[10.2196/jmir.3578](https://doi.org/10.2196/jmir.3578)

KEYWORDS

breast cancer; electronic questionnaires; paper questionnaires; quality of collected data

Introduction

With the evolution of modern technology, electronic applications are increasingly being used in hospitals. Web-based applications and touchscreen devices are finding their way into hospitals for numerous purposes. These electronic applications can be useful

for research purposes, for collecting patient-reported outcomes, and questionnaires [1-3]. Some of the most important advantages of electronic over paper questionnaires include easy usage and immediate electronic storage of results. The use of electronic applications has been evaluated for informed consenting procedures, assessing quality of life, medical education,

interventions, diagnostics, and filling out questionnaires [1,3-11].

Obtaining high response rates without missing information is important for research purposes, as non-responders can bias study results [12]. Response rates have been found to be lower using electronic questionnaires compared to paper questionnaires [13-15]. In order to potentially improve response rates, specific patient subgroups with a preference for electronic questionnaires could be identified. For example, elderly patients may not be as experienced with electronic applications. Aiello et al compared the use of a tablet to paper questionnaires in a mammography clinic. They found that older women (>60 years) had a slightly harder time learning to use the tablet compared to younger patients, but preference towards the tablet was similar in both groups [2].

The aim of our study was to assess differences in demographic characteristics of patients choosing paper versus electronic questionnaires and to evaluate data quality and completeness of data of both approaches.

Methods

Study Context

This study was performed in the University Medical Center Utrecht, the Netherlands (approximate caseload of 180 newly diagnosed breast cancer patients per year). In 2011, same-day diagnosis for breast cancer was introduced with the aim to provide a definitive diagnosis within one day in over 80% of patients. Reducing the time of uncertainty about a diagnosis could potentially reduce anxiety and stress. All patients suspected of breast cancer visited the outpatient breast clinic and underwent physical examination, diagnostic imaging (mammography and ultrasound) with a histological biopsy if indicated, and received a final diagnosis at the end of the day after a multidisciplinary meeting.

Between mid-October 2012 and June 2013, all patients referred to the same-day diagnosis out-patient breast clinic were eligible to participate in the study. Approval for this study was obtained from the local ethics committee, and all patients signed written informed consent. All patients were asked to fill out the 6-item State Trait Anxiety Inventory (STAI) [16,17] questionnaire at six different time points (measuring moments) during the diagnostic phase to evaluate levels of stress and anxiety (Figure 1). Patients were given the opportunity to fill out the

questionnaires electronically (Figure 2) or on paper (Figure 3). Preference towards paper or electronic questionnaires was measured at baseline. The paper questionnaires were returned by mail. In the electronic scenario, the first three questionnaires (administered on the day of diagnosis in the hospital) were offered by means of tablets (iPad). For the last three electronic questionnaires that were to be filled out at home, we used Web-based (hypertext markup language [HTML]) questionnaires. An email with login information to the questionnaires was sent to participants by email on the diagnostic day. The STAI questionnaire was displayed on one page, and all six questions needed to be answered before the form could be submitted. If a question was left blank, an automated message appeared saying that all questions needed to be answered. Patients were not able to look back at previously completed questionnaires. The tablets were also used for providing information and entertainment. An information app was built to provide information on the diagnostic process, diagnostic procedures, treatment team, and routing during the diagnostic day. Several forms of entertainment were available on the tablet, including digital newspapers, magazines, games, and music. The paper questionnaires were returned by mail in a pre-stamped return envelope.

Outcome Measures

Outcome measures included differences in demographic characteristics between patients choosing paper or electronic questionnaires and data quality, focusing on age, reason for referral, breast cancer history, level of education, and baseline anxiety. Data quality was assessed by focusing on missing information, defined as a questionnaire containing at least one unanswered question. To assess if a breast cancer diagnosis affected the quality of data, subgroup analysis including only patients with a benign diagnosis was performed.

Methods for Data Analysis

Demographics, history of breast disease, and diagnostic findings were described as proportions and means with standard deviation. Differences in demographic characteristics, reported anxiety score, and completeness of reported data between the electronic and the paper group were compared by means of chi-square test and independent samples *t* test, where appropriate. Significant differences were defined as *P* values of .05 or less. All statistical analyses were performed using SPSS version 20.0.

Figure 1. Overview of the six measuring moments during the diagnostic phase to evaluate level of stress and anxiety.

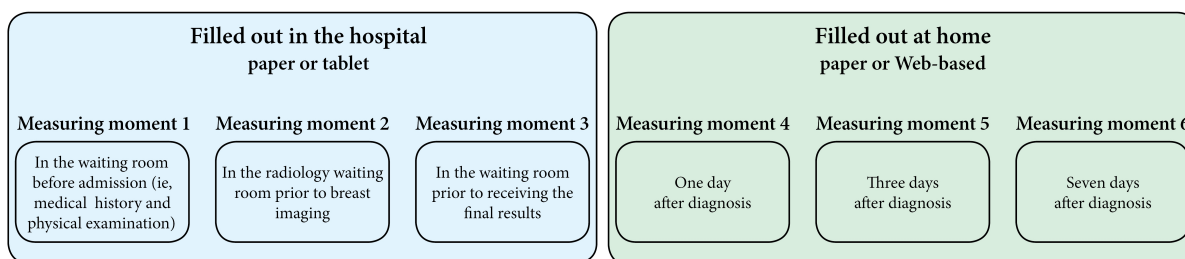


Figure 2. Screenshot of the electronic questionnaire (measuring moment 2).

Snel diagnostiek Mammapoli
Meetmoment 2

18:47

Invulmoment : Tijdens wachttijd voor de onderzoeken bij de Radiologie Reset vragen

Plaats : In de wachtkamer van de radiologie (wachtruimte 5b)

Vragenlijst : **B**

Vragenlijst **B** Hieronder staat een aantal uitspraken die mensen hebben gebruikt om te beschrijven hoe ze zich voelden. We vragen u elke vraag te lezen en de meest passende score aan te klikken om aan te geven hoe u zich **op dit (specifieke) moment** voelt. Er zijn geen goede of foute antwoorden. Denk niet te lang na over de antwoorden, maar klik het antwoord aan dat voor uw gevoel het beste uitdrukt hoe u zich **op dit moment** voelt.

* Dit is een gestandaardiseerde vragenlijst. Sommige vragen kunnen hierdoor op bepaalde momenten als vreemd of niet passend overkomen. We hopen toch dat u steeds alle onderdelen invult. Alleen dan krijgen we een goed beeld van de ervaring en beleving van patiënten tijdens sneldiagnostiek.

Vraag	Helemaal niet	Enigszins	Matig	Heel erg
1. Ik voel me kalm				
2. Ik ben gespannen				
3. Ik ben overstuur				
4. Ik ben ontspannen				
5. Ik voel me tevreden				
6. Ik maak me zorgen				

Toelichting op uw antwoorden

* Vragenlijst B = 6-itemversie van de State Scale van de State-Trait Anxiety Inventory (STAI).

Figure 3. Screenshot of the paper questionnaire (measuring moment 2).



**Universitair Medisch Centrum
Utrecht**

MEETMOMENT 2

S 0050

Invulmoment: tijdens wachttijd voor de onderzoeken bij de Radiologie
Plaats: in de wachtkamer van de Radiologie (wachtruimte 5b).
Vragenlijst: B

Vragenlijst B

Hieronder staat een aantal uitspraken die mensen hebben gebruikt om te beschrijven hoe ze zich voelden. We vragen u elke vraag te lezen en de meest passende score te omcirkelen om aan te geven hoe u zich **op dit (specifieke) moment** voelt. Er zijn geen goede of foute antwoorden. Denk niet te lang na over de antwoorden, maar omcirkel het antwoord dat voor uw gevoel het beste uitdrukt hoe u zich **op dit moment** voelt.

**Dit is een gestandaardiseerde vragenlijst. Sommige vragen kunnen hierdoor op bepaalde momenten als vreemd of niet passend overkomen. We hopen toch dat u steeds alle onderdelen invult. Alleen dan krijgen we een goed beeld van de ervaring en beleving van patiënten tijdens sneldiagnostiek.*

Datum van invullen:.....

	Helemaal niet	Enigszins	Matig	Heel erg
1. Ik voel me kalm	1	2	3	4
2. Ik ben gespannen	1	2	3	4
3. Ik ben overstuur	1	2	3	4
4. Ik ben ontspannen	1	2	3	4
5. Ik voel me tevreden	1	2	3	4
6. Ik maak me zorgen	1	2	3	4

Ruimte voor een toelichting op uw antwoorden:

.....

.....

.....

* Vragenlijst B = 6-itemversie van de State Scale van de State-Trait Anxiety Inventory (STAI)

Results

Demographic Data

Of 321 patients referred to our out-patient breast clinic, 136 patients (42.4%) agreed to participate in the study. All patients were offered the choice of paper or electronic questionnaires.

The mean age was 51.3 years (range 18-85 years) and 35.3% (48/136) patients chose to fill out the questionnaires electronically (Table 1). Reason for referral, family history of breast cancer, and breast-related medical history were similar in both groups. Baseline anxiety scores (as measured by the STAI) did not differ between the groups (46.4 in the paper group versus 43.8 in the electronic group, $P=.30$). Diagnostic imaging findings and proportion of patients undergoing biopsy were similar in both groups. Patients choosing to fill out questionnaires electronically were significantly younger compared to those opting for paper questionnaires (47.3 years

vs 53.5, respectively; $P=.01$) and had a higher level of education ($P=.004$).

Outcome Data: Missing Information

There was significantly more missing information (ie, questionnaires containing at least one unanswered question) in the paper group during the diagnostic day (measuring moments 1-3) compared to the electronic group (28/88 vs 1/48, $P<.001$) (Table 2). In the paper group, this included two patients who did not fill out one or two questions (instead of complete questionnaires not filled out).

In the week after the diagnostic day (measuring moments 4-6), missing information was significantly more prevalent in the electronic group (41/48, 85%) compared to the paper group (38/88, 43%) ($P<.001$). This included 7 patients in the paper group who left one or two questions unanswered. These differences persisted in subgroup analysis including only patients with a benign diagnosis.

Table 1. Demographic characteristics of patients undergoing 1-day diagnosis for suspected breast cancer, comparing patients choosing paper questionnaires (n=88) to those choosing electronic questionnaires (n=48).

Characteristics	Paper, n (%)	Electronic, n (%)	P value ^a
Female	86 (98)	47 (98)	.94 ^b
Age in years, mean (SD), range	53.5 (14.1), 22-85	47.3 (11.4), 18-66	.01 ^c
Reason for referral			
Screening	27 (31)	16 (37)	.75 ^b
Palpable lesion/symptoms	61 (69)	32 (67)	
Positive family history of breast cancer – yes	34 (39)	13 (27)	.18 ^b
Previous breast disease/complaints – yes	28 (32)	19 (40)	.36 ^b
Previous breast cancer diagnosis – yes	4 (5)	3 (6)	.67 ^b
Breast Imaging Reporting and Data System (BI-RADS) classification suspected lesion			
I	18 (21)	9 (19)	.54 ^b
II	34 (39)	22 (46)	
III	9 (10)	9 (19)	
IV	15 (17)	5 (10)	
V	10 (11)	3 (6)	
VI	1 (1)	0 (0)	
No imaging performed	1 (1)	0 (0)	
Biopsy performed			
No	51 (58)	29 (60)	.62 ^b
Yes, histology	32 (36)	18 (38)	
Yes, cytology	5 (6)	1 (2)	
Cancer – yes	16 (18)	7 (15)	.59 ^b
Level of education^d			
Low-moderate ^e	21 (26)	1 (3)	.004 ^b
Moderate-high ^f	27 (33)	10 (29)	
High ^g	33 (41)	24 (69)	
Missing	7	13	
Mean baseline anxiety score, STAI (SD)	46.4 (13.3)	43.8 (13.5)	.30 ^c

^aP values are based on valid proportions.

^bCalculated by chi-square test.

^cCalculated by independent samples *t* test.

^dLevel of education is based on the Dutch educational system

^eLow-moderate education includes primary education/low pre-vocational/secondary general education.

^fModerate-high education includes secondary vocational/higher general and pre-university education.

^gHigh education includes higher vocational education/university.

Table 2. Differences in proportion of patients with incompletely filled out questionnaires between patients opting for paper questionnaires (n=88) and patients choosing electronic questionnaires (n=48).

	Paper, n (%)	Electronic, n (%)	P value
Baseline – moment 1 – STAI 1	25 (28)	0 (0)	
Measuring moment 2 – STAI 2	24 (27)	1 (2)	
Measuring moment 3 – STAI 3	25 (28)	1 (2)	
Measuring moment 4 – STAI 4	28 (32)	27 (56)	
Measuring moment 5 – STAI 5	31 (35)	34 (71)	
Measuring moment 6 – STAI 6	35 (40)	33 (69)	
Measuring moment 1-3 – in hospital ^a	28 (32)	1 (2)	<.001
Measuring moment 4-6 – at home ^b	38 (43)	41 (85)	<.001

^aIncludes all patients with at least 1 incomplete questionnaire in measuring moment 1, 2, or 3.

^bIncludes all patients with at least 1 incomplete questionnaire in measuring moment 4, 5, or 6.

Discussion

Principal Findings

The use of tablets and Web-based questionnaires for collection of patient-reported data has many potential advantages over the use of paper questionnaires. Still, the present study shows that a majority of patients preferred paper over electronic questionnaires. Younger patients and those with a higher level of education were more likely to opt for electronic questionnaires. When using tablets (during the diagnostic day in the hospital), more complete information was collected compared to using paper questionnaires. These data suggests that tablets are superior to paper questionnaires. However, the use of Web-based questionnaires resulted in less complete data collection than paper questionnaires. This might be due to the study design where patients could fill out the electronic questionnaires only on a specific day.

A major advantage of filling out electronic questionnaires is that information is immediately saved. Other advantages of the use of tablets include the possibility of automatically reminding the patient to fill out the questionnaires, and providing information and entertainment. We did not electronically remind patients by email to fill out the questionnaires. Considering the high percentage of incompletely filled out Web-based questionnaires (85%), we would definitively incorporate this in a future study. We did use an automated message when not all questions were answered. This resulted in completely filled out questionnaires in the electronic group, which could possibly lead to more complete data.

Possible drawbacks of using Web-based questionnaires are high non-response rates, impaired reliability and validity, and safety or confidentiality issues [18]. Drawbacks of tablets are the need for upgrades, wireless network unreliability, hardware theft [19], and costs. Fritz et al performed a cost-effectiveness analysis comparing the costs of electronic questionnaires offered on a tablet with paper questionnaires. They found the break-even point to be at 1737 paper sheets per year [1].

Completeness of data collection was very high in the tablet group, with only 1 of the 48 patients not filling out all questions at the first three measuring moments. Missing information was highest in the Web-based group, where many patients (41/48) did not fill out all questions at the last three measuring moments. One likely reason for the high rate of missing information in this group was that patients could fill out the questionnaires only on the correct day (ie, exactly 1, 3, or 7 days after the patient's visit). Our aim was to measure patients' anxiety at specific moments in time, and we limited the possibility of filling out the questionnaires to the correct day only. Patients in the paper group, however, were able to fill out the questionnaires at any given time. This led to more missing information in the Web-based group, and results on missing information need to be interpreted with this information in mind. However, limiting patients to one specific moment to fill out the questionnaires might lead to more accurate measurements of patients' anxiety at that specific moment. For higher response in the Web-based group, automated email reminders could be useful.

Limitations

A possible limitation of this study was that we included only breast cancer patients and consequently, 98% were female. These results are therefore not generalizable to other populations. There could be reduced missing data in the electronic group when other groups are included in a similar study (eg, men, young adults).

Conclusions

Younger patients and patients with a higher level of education have a preference towards filling out questionnaires electronically. In the hospital, a tablet is an excellent medium for patients to fill out questionnaires with very little missing information. However, for filling out questionnaires at home, paper questionnaires result in a better response compared to Web-based questionnaires.

Acknowledgments

The authors thank Peter Sommer and Pieter Ruijssenaars from BeagleBoxx for their cooperation and providing the software and hardware for measuring patient impact of same-day diagnosis.

Conflicts of Interest

None declared.

References

1. Fritz F, Balhorn S, Riek M, Breil B, Dugas M. Qualitative and quantitative evaluation of EHR-integrated mobile patient questionnaires regarding usability and cost-efficiency. *Int J Med Inform* 2012 May;81(5):303-313. [doi: [10.1016/j.ijmedinf.2011.12.008](https://doi.org/10.1016/j.ijmedinf.2011.12.008)] [Medline: [22236957](https://pubmed.ncbi.nlm.nih.gov/22236957/)]
2. Aiello EJ, Taplin S, Reid R, Hobbs M, Seger D, Kamel H, et al. In a randomized controlled trial, patients preferred electronic data collection of breast cancer risk-factor information in a mammography setting. *J Clin Epidemiol* 2006 Jan;59(1):77-81. [doi: [10.1016/j.jclinepi.2005.07.007](https://doi.org/10.1016/j.jclinepi.2005.07.007)] [Medline: [16360564](https://pubmed.ncbi.nlm.nih.gov/16360564/)]
3. Duracinsky M, Lalanne C, Goujard C, Herrmann S, Cheung-Lung C, Brosseau JP, et al. Electronic versus paper-based assessment of health-related quality of life specific to HIV disease: reliability study of the PROQOL-HIV questionnaire. *J Med Internet Res* 2014;16(4):e115 [FREE Full text] [doi: [10.2196/jmir.3330](https://doi.org/10.2196/jmir.3330)] [Medline: [24769643](https://pubmed.ncbi.nlm.nih.gov/24769643/)]
4. Chalil Madathil K, Koikkara R, Obeid J, Greenstein JS, Sanderson IC, Fryar K, et al. An investigation of the efficacy of electronic consenting interfaces of research permissions management system in a hospital setting. *Int J Med Inform* 2013 Sep;82(9):854-863 [FREE Full text] [doi: [10.1016/j.ijmedinf.2013.04.008](https://doi.org/10.1016/j.ijmedinf.2013.04.008)] [Medline: [23757370](https://pubmed.ncbi.nlm.nih.gov/23757370/)]
5. George P, Dumenco L, Doyle R, Dollase R. Incorporating iPads into a preclinical curriculum: a pilot study. *Med Teach* 2013;35(3):226-230. [doi: [10.3109/0142159X.2012.735384](https://doi.org/10.3109/0142159X.2012.735384)] [Medline: [23425119](https://pubmed.ncbi.nlm.nih.gov/23425119/)]
6. Dorr M, Lesmes LA, Lu ZL, Bex PJ. Rapid and reliable assessment of the contrast sensitivity function on an iPad. *Invest Ophthalmol Vis Sci* 2013 Nov;54(12):7266-7273. [doi: [10.1167/iovs.13-11743](https://doi.org/10.1167/iovs.13-11743)] [Medline: [24114545](https://pubmed.ncbi.nlm.nih.gov/24114545/)]
7. Glaser D, Jain S, Kortum P. Benefits of a physician-facing tablet presentation of patient symptom data: comparing paper and electronic formats. *BMC Med Inform Decis Mak* 2013;13:99 [FREE Full text] [doi: [10.1186/1472-6947-13-99](https://doi.org/10.1186/1472-6947-13-99)] [Medline: [24004844](https://pubmed.ncbi.nlm.nih.gov/24004844/)]
8. Mayfield CH, Ohara PT, O'Sullivan PS. Perceptions of a mobile technology on learning strategies in the anatomy laboratory. *Anat Sci Educ* 2013;6(2):81-89. [doi: [10.1002/ase.1307](https://doi.org/10.1002/ase.1307)] [Medline: [22927203](https://pubmed.ncbi.nlm.nih.gov/22927203/)]
9. Geraedts AS, Kleiboer AM, Twisk J, Wiezer NM, van Mechelen W, Cuijpers P. Long-term results of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. *J Med Internet Res* 2014;16(7):e168 [FREE Full text] [doi: [10.2196/jmir.3539](https://doi.org/10.2196/jmir.3539)] [Medline: [25008127](https://pubmed.ncbi.nlm.nih.gov/25008127/)]
10. Lalloo C, Kumbhare D, Stinson JN, Henry JL. Pain-QuILT: clinical feasibility of a web-based visual pain assessment tool in adults with chronic pain. *J Med Internet Res* 2014;16(5):e127 [FREE Full text] [doi: [10.2196/jmir.3292](https://doi.org/10.2196/jmir.3292)] [Medline: [24819478](https://pubmed.ncbi.nlm.nih.gov/24819478/)]
11. Hanberger L, Ludvigsson J, Nordfeldt S. Use of a web 2.0 portal to improve education and communication in young patients with families: randomized controlled trial. *J Med Internet Res* 2013;15(8):e175 [FREE Full text] [doi: [10.2196/jmir.2425](https://doi.org/10.2196/jmir.2425)] [Medline: [23973555](https://pubmed.ncbi.nlm.nih.gov/23973555/)]
12. Partin MR, Malone M, Winnett M, Slater J, Bar-Cohen A, Caplan L. The impact of survey nonresponse bias on conclusions drawn from a mammography intervention trial. *J Clin Epidemiol* 2003 Sep;56(9):867-873. [Medline: [14505772](https://pubmed.ncbi.nlm.nih.gov/14505772/)]
13. Hohwü L, Lyshol H, Gissler M, Jonsson SH, Petzold M, Obel C. Web-based versus traditional paper questionnaires: a mixed-mode survey with a Nordic perspective. *J Med Internet Res* 2013;15(8):e173 [FREE Full text] [doi: [10.2196/jmir.2595](https://doi.org/10.2196/jmir.2595)] [Medline: [23978658](https://pubmed.ncbi.nlm.nih.gov/23978658/)]
14. Kongsved SM, Basnov M, Holm-Christensen K, Hjollund NH. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. *J Med Internet Res* 2007;9(3):e25 [FREE Full text] [doi: [10.2196/jmir.9.3.e25](https://doi.org/10.2196/jmir.9.3.e25)] [Medline: [17942387](https://pubmed.ncbi.nlm.nih.gov/17942387/)]
15. Leece P, Bhandari M, Sprague S, Swiontkowski MF, Schemitsch EH, Tornetta P, et al. Internet versus mailed questionnaires: a controlled comparison. *J Med Internet Res* 2004 Oct 29;6(4):e39 [FREE Full text] [doi: [10.2196/jmir.6.4.e39](https://doi.org/10.2196/jmir.6.4.e39)] [Medline: [15631963](https://pubmed.ncbi.nlm.nih.gov/15631963/)]
16. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br J Clin Psychol* 1992 Sep;31 (Pt 3):301-306. [Medline: [1393159](https://pubmed.ncbi.nlm.nih.gov/1393159/)]
17. van der Bij AK, de Weerd S, Cikot RJ, Steegers EA, Braspenning JC. Validation of the Dutch short form of the state scale of the Spielberger State-Trait Anxiety Inventory: considerations for usage in screening outcomes. *Community Genet* 2003;6(2):84-87. [Medline: [14560068](https://pubmed.ncbi.nlm.nih.gov/14560068/)]
18. 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](https://doi.org/10.1093/aje/kwq291)] [Medline: [20880962](https://pubmed.ncbi.nlm.nih.gov/20880962/)]

19. McDonnell CH, Seidenwurm DJ, McDonnell DE, Bobolis KA. Self administered screening for hereditary cancers in conjunction with mammography and ultrasound. *Fam Cancer* 2013 Dec;12(4):651-656. [doi: [10.1007/s10689-013-9641-z](https://doi.org/10.1007/s10689-013-9641-z)] [Medline: [23584879](https://pubmed.ncbi.nlm.nih.gov/23584879/)]

Abbreviations

BI-RADS: Breast Imaging Reporting and Data System

HTML: hypertext markup language

STAI: State Trait Anxiety Inventory

Edited by G Eysenbach; submitted 03.06.14; peer-reviewed by F Fritz, L Hohwü, C Schäfer; comments to author 17.07.14; revised version received 14.08.14; accepted 23.08.14; published 31.10.14.

Please cite as:

Barentsz MW, Wessels H, van Diest PJ, Pijnappel RM, Haaring C, van der Pol CC, Witkamp AJ, van den Bosch MA, Verkooijen HM. Tablet, Web-Based, or Paper Questionnaires for Measuring Anxiety in Patients Suspected of Breast Cancer: Patients' Preferences and Quality of Collected Data

J Med Internet Res 2014;16(10):e239

URL: <http://www.jmir.org/2014/10/e239/>

doi: [10.2196/jmir.3578](https://doi.org/10.2196/jmir.3578)

PMID: [25364951](https://pubmed.ncbi.nlm.nih.gov/25364951/)

©Maarten W Barentsz, Hester Wessels, Paul J van Diest, Ruud M Pijnappel, Cees Haaring, Carmen C van der Pol, Arjen J Witkamp, Maurice A van den Bosch, Helena M Verkooijen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 31.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Digital Inequalities of Family Life Information Seeking and Family Well-Being Among Chinese Adults in Hong Kong: A Population Survey

Man Ping Wang^{1*}, PhD; Xin Wang^{2*}, MSc; Kasisomayajula Viswanath^{3,4*}, PhD; Alice Wan^{2*}, MSc; Tai Hing Lam^{2*}, MD; Sophia S Chan^{1*}, PhD

¹School of Nursing, The University of Hong Kong, Hong Kong, China (Hong Kong)

²School of Public Health, The University of Hong Kong, Hong Kong, China (Hong Kong)

³Department of Social and Behavioral Sciences, Harvard School of Public Health, Boston, MA, United States

⁴Center for Community-Based Research, Dana-Farber Cancer Institute, Boston, MA, United States

* all authors contributed equally

Corresponding Author:

Tai Hing Lam, MD
School of Public Health
The University of Hong Kong
5/F WMMW Block
23 Sassoon Rd., Pokfulam
Hong Kong,
China (Hong Kong)
Phone: 852 28199287
Fax: 852 28559528
Email: hrrmlth@hku.hk

Abstract

Background: Inequalities in Internet use and health information seeking are well documented, but less is known about information for family life activities.

Objective: We investigated the social determinants of online family life information seeking behaviors and its associations with family well-being among Chinese adults in Hong Kong.

Methods: A probability-based telephone survey was conducted in 2012 to record family life information seeking behaviors, including frequency of seeking and paying attention to family life information, levels of trust, and perceived usefulness of family life information. Family well-being was assessed using 3 single items on perceived family harmony, happiness, and health, with higher scores indicating greater well-being. Adjusted odds ratios for family life information seeking behaviors by socioeconomic characteristics and lifestyle behaviors, and adjusted beta coefficients for family well-being by family life information seeking behaviors were calculated.

Results: Of 1537 respondents, 57.57% (855/1537) had ever and 26.45% (407/1537) sought monthly family life information through the Internet. Lower educational attainment and household income, smoking, and physical inactivity were associated with less frequent seeking and paying attention (all $P < .05$). Greater perceived family health was associated with more frequent attention (adjusted $\beta = .32$, 95% CI .11-.52), greater levels of trust (adjusted $\beta = .28$, 95% CI .07-.48), and perceived usefulness (adjusted $\beta = .23$, 95% CI .01-.45) of family life information. Frequent attention and higher level of trust were also associated with greater family harmony (adjusted $\beta = .22$, 95% CI .002-.41) and happiness (adjusted $\beta = .23$, 95% CI .003-.42), respectively.

Conclusions: This is the first study investigating family life information seeking behaviors and suggested inequalities of online family life information seeking behaviors. The association between family life information seeking behavior and family well-being needs to be confirmed in prospective studies.

(*J Med Internet Res* 2014;16(10):e227) doi:[10.2196/jmir.3386](https://doi.org/10.2196/jmir.3386)

KEYWORDS

socioeconomic factors; information seeking behavior; family health

Introduction

Communication inequalities, defined as differences in accessing, processing, and acting on information, may be the link between social determinants and health [1]. Diffusion of advanced information and communication technologies (ICTs) in Hong Kong has led to a high rate of Internet penetration (77.9%), but a digital divide (defined as differential access to Internet among different groups of people) [2] between socioeconomic classes [3]. Prevalence of Internet connection among people living in public housing estates (69.3%) and having low monthly household income (<HK\$10,000, 38.9%) was much lower than private housing (82.1%) and income >HK\$50,000 (98.2%) [3]. A similar phenomenon was observed for computer use and educational gradients (primary or less: 28.4%; secondary: 82.3%; and postsecondary or greater: 96.8%) [3].

Such inequalities on Internet use suggest a widening gap between information haves and have nots as the Internet is increasingly used as a source for information seeking, although offline searching (eg, television and newspaper) remains an important channel for obtaining information. Bridging communication inequalities is pivotal to reduce social inequality [4]. Most studies have focused on digital health communication inequalities and adverse effects across the continuum of health [5-8]. Inequalities in health information communication only represent a part of the digital divide because the Internet is also used for everyday life information seeking (ELIS), which includes information for solving problems encountered in everyday situations [9,10]. ELIS is studied mostly in the field of information science and rarely in behavioral or social science. Therefore, we aimed to extend current research on digital divide to include information on family life activities.

Family life information is considered an important part of family life, and refers to information to strengthen family functioning through improving communication skills, knowledge about developmental tasks, decision-making skills, self-esteem, and interpersonal relationships [11]. Only a few studies have investigated the family information needs for specific topics, such as parenting, childcare, and information needs for sick children [12-14]. Recognizing the importance of providing family life information, particularly for parents with young children, the UK government has set up family information services in all parts of the country to provide comprehensive family-related information particularly for lower socioeconomic groups [15].

The concept of family life information has not been clearly defined. We adopted a broad, simple, and practical definition of family life information seeking behaviors as information related to family communication; relationships with children, partners, and other family members; work-life balance; and emotion and stress management. These components are commonly reported in many Western studies as main factors affecting family well-being [16,17]. Similarly, our qualitative studies exploring the concepts of family well-being in the general public and community leaders in Hong Kong found that health, happiness, and harmony (3Hs) are the 3 major themes of family well-being [18,19]. The concept of the family 3Hs

are coherent with the traditional Chinese value on collectivism, but are different from Western individualist culture, which puts more emphasis on personal happiness and independence than family 3Hs [20].

Specifically, this study aimed to investigate socioeconomic and sociodemographic correlates of frequency of seeking, frequency of attention, levels of trust, and perceived usefulness toward online family life information focusing on socioeconomic inequalities. In addition, the associations between family life information seeking behaviors and the family 3Hs were also investigated.

Methods

Sampling

Details of research design have been reported elsewhere [6,21]. In brief, as a part of the FAMILY project, the Hong Kong Family and Health Information Trends Survey (FHInTS) was conducted in 2012 using a random telephone-based survey of the general public to monitor opinions and behaviors related to family health and communication [6]. All interviews were conducted by trained interviewers of the Public Opinion Programme from the University of Hong Kong. A 2-stage random sampling method was used. Telephone numbers were retrieved from residential telephone directories that cover approximately 76% of Hong Kong residents [22]. A computer program was used to generate a list of the telephone numbers in random order for interviews. Invalid household numbers, nonresponses, and ineligible households (people aged <18 years or not able to speak Cantonese) were excluded (n=8748). In the second stage, after interviewers introduced the study purpose, adult respondents were asked how many eligible persons were living in the household. All eligible persons were listed and the person with the next birthday closest to the interview day was selected for interview. Each interview took approximately 25 minutes to complete. Among 2080 people with confirmed eligibility, 1537 adults were successfully interviewed yielding a response rate of 73.9% [23]. Ethical approval was granted by the Institutional Review Board (IRB) of the University of Hong Kong/Hospital Authority Hong Kong West Cluster. Verbal informed consent was obtained and recorded verbatim, and the procedure was approved by the IRB.

Measurement

Definitions of families (family members are those who have relationships through biological, marital, cohabital, or emotional bonding) and family life information seeking (definition as mentioned previously) were presented to the respondents prior to the questions about family life information seeking. Frequency of family life information seeking was assessed by the question: "In the past 12 months, how often have you searched for family life information on the Internet?" with responses of ≥ 1 time/week, 1-3 times/month, 1 time in several months, rarely, and never/do not have Internet access. Attention to family information was assessed by the question "How often did you pay attention to family life information" with responses of frequently, sometimes, rarely, and never. Trust in family life information was assessed by the question: "What do you think about the online family information sought last time?" with

responses of very trustful, partly trustful, neutral, not trustful, and completely not trustful. Perceived usefulness of online family life information was assessed by the question: "Do you think online family life information is useful?" with responses of very useful, partly useful, little useful, and not useful. Perceived family harmony, happiness, and health was assessed using 3 separate questions asking respondents to give a score from 0-10 with a higher score indicating better family well-being. Internal consistency of these 3 items was supported by a satisfactory Cronbach alpha (.84).

As in other similar studies on information seeking [5,24], socioeconomic status (SES) was measured using educational attainment, household monthly income, and employment. Several studies have documented the influence of these SES variables on a variety of health outcomes [25-27]. The responses for SES were based on our previous studies and the Hong Kong census with slight modifications. Education attainment was categorized as primary or less (combining no formal education and primary education), secondary, and tertiary or greater. Monthly household income (in Hong Kong dollars; US \$1 = HK\$7.8) was categorized as <HK\$10,000 (combining <HK\$4000 and HK\$4000-HK\$9999), HK\$10,000-HK\$19,999, HK\$20,000-HK\$29,999, HK\$30,000-HK\$39,999, and ≥HK\$40,000. Employment status was categorized as full-time, part-time, self-employed, and unemployed. Health behaviors were correlated with information seeking as found in our previous studies [6]. We assessed the associations of smoking (never, ex-, and current), alcohol drinking (never, ex-, occasional, and monthly), and physical activity (none, 1-3 days/week, ≥4 days/week) with family life information seeking.

Statistical Analysis

All data were weighted by sex and age from the census data. Inequalities in family life information seeking by sex, age, marital status, and SES indicators were assessed by logistic

regression, which yielded adjusted odds ratios (aOR) for family life information seeking. The association between family life information seeking and family well-being was estimated using linear regression (beta coefficients) adjusted for sociodemographic characteristics and health status. All analyses were performed using STATA 10 (StataCorp LP, College Station, TX, USA). A *P* value of <.05 was considered statistically significant.

Results

Of all respondents (N=1537), 45.71% (703/1537) were male, 73.78% (1123/1522) aged 25-64 years, 61.13% (937/1533) were married, 86.59% (1325/1530) had secondary school educational attainment, and 61.81% (801/1296) had monthly household income ≥HK\$ 20,000. Details of socioeconomic status are reported elsewhere [28]. Sample representativeness was supported by small difference in distribution of sex, age, educational attainment, and household income between our sample and the general population (Cohen's effect size <0.3) [28].

Table 1 shows that one-quarter of respondents (25.8%, 396/1537) sought family life information for recreational purposes followed by information for improving family relationships (17.7%, 272/1537) and communication (15.3%, 235/1537), managing emotional problems and stress (14.0%, 215/1537), and improving work and ability (10.8%, 165/1537). More than half of the respondents (57.57%, 885/1537) had ever sought family life information with 26.45% (407/1537) on a monthly basis. In addition to active seeking, more than two-thirds (69.12%, 1059/1532) of the respondents had ever paid attention to family life information. Only 3.2% (39/1198) reported online family life information as trustful and 10.5% (137/1309) of respondents reported it as very useful.

Table 1. Online family life information seeking content, frequency, attention, and trust (N=1537).

Content	Unweighted, n (%)	Weighted, n (%)
Entertainment information	334 (21.73)	396 (25.75)
Family relationship	223 (14.51)	272 (17.68)
Children	177 (11.52)	208 (13.53)
Couples	104 (6.77)	135 (8.77)
Relatives	28 (1.82)	37 (2.42)
Family communication	199 (12.95)	235 (15.27)
Emotion and stress management	178 (11.58)	215 (14.02)
Improvement of self/work-ability	134 (8.72)	165 (10.72)
Work-life balance	66 (4.29)	87 (5.67)
Ability of self-independent	59 (3.84)	77 (5.02)
Others	42 (2.73)	45 (2.93)
Frequency of seeking		
≥1 time(s)/week	98 (6.38)	120 (7.78)
1-3 times/month	242 (15.76)	287 (18.67)
<1 time/month	423 (27.54)	478 (31.12)
Never	773 (50.33)	652 (42.43)
Attention		
Always	189 (12.37)	228 (14.87)
Sometimes	438 (28.66)	491 (32.06)
Seldom	316 (20.68)	340 (22.19)
Never	585 (38.29)	473 (30.88)
Trust		
Trust	38 (3.50)	39 (3.24)
Some trust	587 (54.00)	657 (54.88)
Neutral	291 (26.77)	328 (27.38)
Not trust	150 (13.80)	152 (12.66)
Absolutely not trust	21 (1.93)	22 (1.83)
Perceived information usefulness		
Very useful	122 (9.78)	137 (10.48)
Partly useful	625 (50.08)	714 (54.56)
Little useful	182 (14.58)	186 (14.20)
Not useful	319 (25.56)	272 (20.77)

Table 2 shows that age was inversely associated with monthly seeking, frequent attention, and perceived usefulness toward family life information (P for trend $<.001$) although no association was observed for trust. Educational attainment was strongly associated with seeking, attention, trust, and perceived usefulness of family life information (all P for trend $<.008$ - $<.001$). Similarly, compared with family income $<HK\$10,000$, higher household income was generally associated

with greater odds of seeking, attention, and perceived usefulness toward family life information although nonsignificant associations for the highest income category of $\geq HK\$30,000$ with attention and perceived usefulness were found. Marital and employment statuses were not associated with family life information seeking behaviors. In addition, family income was significantly associated with perceived levels of family health ($\beta=.37$, $P=.04$) and harmony ($\beta=.90$, $P=.005$).

Table 2. Socioeconomic status and online family life information seeking behaviors (N=1537).

Factors	Monthly seeking		Frequent attention		Trust		Perceived usefulness	
	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a
Sex								
Male	198 (28.2)	1	342 (48.9)	1	319 (56.0)	1	407 (66.9)	1
Female	208 (25.0)	0.90 (0.69-1.19)	376 (45.3)	0.92 (0.71-1.20)	376 (60.1)	1.12 (0.85-1.47)	444 (63.4)	0.76 (0.56-1.02)
Age (years)								
18-24	68 (43.8)	1	102 (65.4)	1	96 (61.8)	1	115 (74.5)	1
25-44	212 (37.8)	0.69 (0.43-1.12)	361 (64.4)	0.77 (0.46-1.27)	323 (59.4)	1.02 (0.63-1.64)	411 (76.3)	1.40 (0.80-2.46)
45-64	113 (20.2)	0.38 (0.22-0.65) ^d	225 (40.4)	0.42 (0.24-0.72) ^c	229 (55.3)	0.92 (0.54-1.56)	273 (59.1)	0.85 (0.46-1.57)
≥65	10 (4.3)	0.09 (0.04-0.25) ^d	26 (10.9)	0.12 (0.06-0.25) ^d	39 (54.3)	1.47 (0.65-3.31)	45 (31.3)	0.28 (0.13-0.60) ^c
<i>P</i> for trend	<0.0001		<0.0001		0.89		<0.0001	
Marital status								
Single	185 (38.0)	1		1		1		1
Married/cohabitated	214 (22.7)	1.02 (0.73-1.43)	308 (63.2)	1.03 (0.74-1.45)	277 (59.6)	1.03 (0.73-1.44)	354 (76.2)	0.77 (0.51-1.14)
Other	7 (7.4)	0.75 (0.28-1.99)	394 (41.9)	0.67 (0.31-1.46)	399 (57.8)	0.81 (0.36-1.84)	471 (60.6)	0.96 (0.44-2.08)
Employment status								
			16 (15.7)		17 (44.3)		26 (40.4)	
Full-time	252 (35.2)	1		1		1		1
Part-time	32 (27.7)	1.18 (0.71-1.97)	425 (59.4)	0.73 (0.45-1.19)	395 (60.5)	0.87 (0.52-1.44)	492 (73.8)	0.95 (0.56-1.61)
Self-employed	16 (22.1)	0.70 (0.38-1.29)	50 (43.6)	0.79 (0.46-1.37)	51 (54.3)	0.65 (0.38-1.13)	65 (64.3)	0.62 (0.35-1.11)
Unemployed	107 (16.8)	0.93 (0.65-1.33)	34 (46.3)	0.88 (0.63-1.23)	31 (50.0)	1.07 (0.75-1.52)	39 (59.3)	1.16 (0.80-1.68)
Education								
			210 (33.4)		218 (56.4)		254 (53.7)	
≤Primary	3 (1.6)	1		1		1		1
Secondary	170 (23.7)	7.42 (2.38-23.12) ^d	12 (6.0)	5.44 (2.81-10.52) ^d	25 (39.0)	2.23 (1.21-4.12) ^b	26 (22.1)	2.98 (1.72-5.16) ^d
≥Tertiary	232 (38.2)	11.04 (3.47-35.12) ^d	305 (42.8)	10.69 (5.34-21.43) ^d	310 (55.5)	2.75 (1.44-5.25) ^c	368 (60.6)	6.30 (3.44-11.54) ^d
<i>P</i> for trend	<0.0001		397 (65.5)		<0.0001		357 (78.9)	
Household income (HK\$)								
≤9999	16 (7.5)	1		1		1		1
10,000-19,999	78 (28.2)	2.02 (1.09-3.76) ^b	39 (17.8)	1.84 (1.13-2.99) ^b	44 (50.0)	1.26 (0.73-2.16)	56 (40.2)	1.40 (0.85-2.32)

Factors	Monthly seeking		Frequent attention		Trust		Perceived usefulness	
	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a
20,000-29,999	74 (28.8)	1.94 (1.03-3.64) ^b	134 (49.3)	1.84 (1.12-3.02) ^b	122 (54.5)	1.48 (0.85-2.57)	152 (63.6)	1.95 (1.15-3.31) ^b
≥30,000	194 (35.7)	2.10 (1.14-3.88) ^b	134 (52.6)	1.58 (0.98-2.56)	127 (59.3)	1.55 (0.90-2.65)	162 (72.3)	1.46 (0.88-2.42)
<i>P</i> for trend		0.11		0.42		0.10		0.24

^a aOR: adjusted odds ratio; mutually adjusted for the variables in the table.

^b $P < .05$.

^c $P < .01$.

^d $P < .001$.

Compared with never smoking, current smoking was associated with lower aORs of 0.35 (95% CI 0.19-0.64) and 0.58 (95% CI 0.35-0.96) for monthly seeking and perceived usefulness of family life information, respectively (Table 3). Greater level of moderate physical activity (≥ 4 days per week) was associated with family life information seeking behaviors particularly for monthly seeking (aOR 1.66, 95% CI 1.19-2.32), frequent attention (aOR 1.69, 95% CI 1.64-2.31), and trust (aOR 1.43, 95% CI 1.04-1.96) of family life information.

Greater level of perceived family health was significantly associated with frequent attention (adjusted $\beta = .32$, 95% CI .11-.52), trust (adjusted $\beta = .28$, 95% CI .07-.48), and perceived usefulness (adjusted $\beta = .23$, 95% CI .01-.45) of family life information, and marginally associated with monthly seeking of family life information (adjusted $\beta = .15$, 95% CI -.06 to .36) (Table 4). In contrast, only frequent attention was associated with family harmony (adjusted $\beta = .22$, 95% CI .02-.41) and trust of the information was associated with family happiness (adjusted $\beta = .23$, 95% CI .03-.42).

Table 3. Behavioral correlates and family life information seeking behaviors (N=1537).

Health behaviors	Monthly seeking		Frequent attention		Trust		Perceived usefulness	
	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a	n (%)	aOR (95% CI) ^a
Smoking								
Never	367 (28.7)	1	612 (48.0)	1	604 (59.6)	1	743 (67.2)	1
Ex-smoker	20 (17.8)	0.78 (0.43-1.42)	39 (34.7)	0.97 (0.57-1.64)	42 (59.2)	0.95 (0.55-1.66)	42 (51.3)	0.58 (0.33-1.02)
Current smoker	17 (11.9)	0.35 (0.19-0.64) ^d	66 (47.5)	1.40 (0.87-2.24)	49 (44.7)	0.72 (0.44-1.17)	63 (53.4)	0.58 (0.35-0.96) ^b
Drinking								
Never	175 (23.4)	1	305 (41.0)	1	310 (57.9)	1	379 (61.7)	1
Ex-drinker	6 (18.4)	2.30 (0.73-7.21)	8 (27.7)	1.17 (0.41-3.37)	10 (57.8)	0.98 (0.33-2.88)	9 (44.8)	1.40 (0.44-4.48)
Occasional drinker	64 (26.8)	1.04 (0.71-1.54)	112 (46.7)	0.88 (0.61-1.27)	115 (58.0)	0.93 (0.64-1.35)	139 (65.2)	0.84 (0.56-1.26)
Monthly drinker	162 (31.3)	1.08 (0.79-1.46)	293 (56.7)	1.06 (0.79-1.42)	260 (58.5)	0.98 (0.73-1.31)	324 (70.4)	0.98 (0.70-1.36)
Moderate physical activity								
None	158 (23.5)	1	281 (42.0)	1	272 (53.1)	1	359 (62.5)	1
1-3 days/week	116 (30.3)	1.30 (0.94-1.80)	217 (56.6)	1.47 (1.08-2.01) ^b	208 (64.6)	1.53 (1.12-2.11) ^c	251 (73.6)	1.47 (1.03-2.11) ^b
≥4 days/week	130 (27.3)	1.66 (1.19-2.32) ^c	219 (46.1)	1.69 (1.64-2.31) ^d	214 (59.5)	1.43 (1.04-1.96) ^b	239 (61.4)	1.05 (0.75-1.46)

^a aOR: adjusted odds ratio; adjusted for sex, age, marital status, employment, education, household income, and diseases.

^b $P < .05$.

^c $P < .01$.

^d $P < .001$.

Table 4. Online family life information seeking behaviors and family well-being (N=1537).

Family information seeking	Family harmony		Family happiness		Family health	
	Mean (SD)	β (95% CI) ^a	Mean (SD)	β (95% CI) ^a	Mean (SD)	β (95% CI) ^a
Monthly seeking						
No	7.6 (1.6)	0	7.4 (1.6)	0	7.2 (1.7)	0
Yes	7.6 (1.6)	.15 (-.06, .36)	7.4 (1.5)	.04 (-.16, .24)	7.3 (1.6)	.14 (-.07, .35)
Frequent attention						
No	7.6 (1.6)	0	7.3 (1.6)	0	7.1 (1.7)	0
Yes	7.7 (1.7)	.22 (.02, .41) ^b	7.4 (1.6)	.14 (-.06, .33)	7.4 (1.6)	.32 (.11, .52) ^c
Trust						
No	7.4 (1.6)	0	7.2 (1.6)	0	7.0 (1.7)	0
Yes	7.6 (1.7)	.14 (-.07, .34)	7.5 (1.5)	.23 (.03, .42) ^b	7.3 (1.6)	.28 (.07, .48) ^c
Perceived usefulness						
No	7.6 (1.7)	0	7.4 (1.7)	0	7.1 (1.8)	0
Yes	7.6 (1.6)	.14 (-.09, .36)	7.4 (1.5)	.02 (-.20, .23)	7.3 (1.6)	.23 (.01, .45) ^b

^a Adjusted for sex, age, marital status, employment, education, household income, and diseases.

^b $P < .05$.

^c $P < .01$.

Discussion

This is probably the first study investigating online family life information seeking behaviors including frequency of seeking and attention, level of trust, and perceived usefulness. In our study, prevalence of online family life information seeking (57.6%) was higher than online health information seeking (44.0%) in Hong Kong [29], but was comparable to health information seeking (72%) in the United States [30], suggesting studying family life information seeking behavior is important. Our study has extended the findings of previous studies on parenting and childcare information [13,14], and revealed that a wide range of family life topics were sought through the Internet, including family recreation, communication, and work-life balance. As a part of family life education in Hong Kong [31], basic online family life information is provided by the Social Welfare Department, Education Bureau, and some nongovernmental organizations (NGOs). On the other hand, the private sector generated numerous online platforms for sharing comprehensive and vibrant family life information in an interactive way. However, the quality and accuracy of the information from these websites is unknown.

Although many respondents paid attention to online family life information (69.1%), very few (3.2%) trusted the information, and only some (10.5%) perceived the information as very useful. The quality of online information is always a concern and similar to patterns seen with health information [32]. This suggests that online platforms providing evidence-based and comprehensive family life information are needed in Hong Kong and elsewhere. Online family life information can be provided by large government sectors or reputable NGOs. Similar websites have been established by the UK National Association of Family Information Services as a mandatory service for

providing family life information [15]. Western studies have shown that websites providing family life education were well received by various groups of people [33]. Incorporating family life information into social media and mobile Internet devices will substantially increase the penetration and adoption of such information. Providing the information will not only benefit Hong Kong Chinese, but also millions of families in mainland China and beyond.

We found that online family life information seeking behavior was socially patterned with lower levels of educational attainment and income associated with less frequent family life information seeking and attention. The results were consistent with the digital divide of health information seeking in local and international studies [5,6,24]. Lack of Internet access, lower social support, and poor information literacy and skills among disadvantaged groups are documented barriers for seeking online information [34]. Compared with household income, educational attainment was an even stronger factor associating with information seeking, level of trust, and perceived usefulness. This suggests that cognitive skills are more influential on family life information seeking behaviors than physical access to the Internet. Unlike health information, the concept of family life information is more vague and less developed; thus, high cognitive functions are needed to refine searching, understand sophisticated content, and translate information into behaviors. Online family life information seeking was also patterned by age and lifestyle behaviors. Increasing age, current smoking status, and physically inactivity were risk factors for infrequent family life information seeking and lower level of perceived usefulness. Our findings can help campaign planners of family life information seeking to focus resources on these groups.

More importantly, we found that family life information seeking was associated with greater family well-being, particularly

perceived family health and harmony, which were also patterned by socioeconomic status. Although our results need to be confirmed in prospective studies, the findings were in-line with benefits of family life education in preventing family problems and improving family functions [35]. Maintaining family well-being is a challenge particularly for families with long working hours, which is typical in Hong Kong (average weekly working 45 hours) [36] comparable to South Korea (44.6 hours) having the longest working hours in countries of Organisation for Economic Co-operation and Development (OECD). Seeking online family life information is convenient and would help these busy families to better prepare for family activities, effectively solve family problems, and demonstrate care of other family members. Future qualitative and quantitative studies are warranted to investigate the underlying mechanisms between family life information seeking and family well-being.

This study has several limitations. First, a cross-sectional design was used and temporality of the associations cannot be confirmed. It is unlikely that family life information seeking would lead to higher educational attainment and income. Nevertheless, greater levels of family well-being may facilitate seeking of family life information. Prospective studies are needed to confirm the associations and test the mediation effects of family life information seeking on the association between socioeconomic status and the family 3Hs. Second, to the best of our knowledge this is the first study to describe the family life information seeking behaviors. Although a simple description of family life information seeking has been provided,

we are uncertain about the variation of family life information seeking definitions perceived by the respondents and how this would affect the results. Future studies are needed to better define family life information seeking probably using a qualitative research design and purposive sampling surveys among respondents who may take a more active role in certain activities in the family to understand the detail of family life information seeking behaviors. The family 3Hs was used a proxy of family well-being as supported by our previous qualitative studies [18,19]. Data from another survey also in the FAMILY project [37] showed family 3Hs items were moderately correlated (Pearson r ranges .34-.47) with family functioning (APGAR) [38], and resilience (Family Resilience Assessment Scale). However, given the distinct difference of perceptions on family functions between cultures, we are not certain about the generalizability of the association between family life information seeking and the family 3Hs to other countries. Third, although our sample was representative of the general population, the effects of nonresponse bias and problems of decreasing landline telephone coverage on the observed associations were uncertain.

This study is the first to investigate family life information seeking behaviors. The results showed that people with lower SES were less likely to seek and pay attention to online family life information or perceive family life information as useful. The associations of family life information seeking with perceived family health and harmony need to be confirmed by prospective studies.

Acknowledgments

The project was funded by the Hong Kong Jockey Club Charities Trust. We would like to thank the participants who participated in the telephone surveys and Public Opinion Programme (HKU) for conducting the surveys.

Conflicts of Interest

None declared.

References

1. Viswanath K. Public communications role in reducing and eliminating health disparities. In: Thomson GE, Mitchell F, Williams MB, editors. Examining the Health Disparities Research Plan of the National Institute of Health: Unfinished Business. Washington, DC: Institute of Medicine; 2006.
2. US Department of Commerce, National Telecommunications and Information Administration. Falling through the Net: Defining the Digital Divide. Washington, DC: US Department of Commerce; 1999.
3. Census Statistics Department. Hong Kong monthly digest of statistics feature article: Usage of personal computers & Internet services by Hong Kong residents 2000 to 2012. In: Hong Kong SAR Government. Hong Kong SAR: Hong Kong SAR Government; 2013:1-13.
4. Warschauer M. Technology and Social Inclusion: Rethinking the Digital Divide. Cambridge, MA: MIT Press; 2014.
5. Ishikawa Y, Nishiuchi H, Hayashi H, Viswanath K. Socioeconomic status and health communication inequalities in Japan: a nationwide cross-sectional survey. *PLoS One* 2012;7(7):e40664 [FREE Full text] [doi: [10.1371/journal.pone.0040664](https://doi.org/10.1371/journal.pone.0040664)] [Medline: [22808229](https://pubmed.ncbi.nlm.nih.gov/22808229/)]
6. Wang MP, Viswanath K, Lam TH, Wang X, Chan SS. Social determinants of health information seeking among Chinese adults in Hong Kong. *PLoS One* 2013;8(8):e73049 [FREE Full text] [doi: [10.1371/journal.pone.0073049](https://doi.org/10.1371/journal.pone.0073049)] [Medline: [24009729](https://pubmed.ncbi.nlm.nih.gov/24009729/)]
7. 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](https://doi.org/10.1177/1460458211414843)] [Medline: [21937464](https://pubmed.ncbi.nlm.nih.gov/21937464/)]

8. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the Internet, and the digital divide. *Health Aff (Millwood)* 2000;19(6):255-265 [FREE Full text] [Medline: [11192412](#)]
9. Savolainen R. Everyday life information seeking: approaching information seeking in the context of way of life. *Libr Inform Sci Res* 1995;17:259-294.
10. Savolainen R. The role of the internet in information seeking. Putting the networked services in context. *Inform Process Manag* 1999;35:765-782.
11. The National Council on Family Relations. 2013. What is Family Life Education URL: <http://www.ncfr.org/cfle-certification/what-family-life-education> [accessed 2014-03-07] [WebCite Cache ID 6NtAdt0Ft]
12. Wakefield CE, Butow P, Fleming CA, Daniel G, Cohn RJ. Family information needs at childhood cancer treatment completion. *Pediatr Blood Cancer* 2012 Apr;58(4):621-626. [doi: [10.1002/pbc.23316](#)] [Medline: [21910212](#)]
13. Devolin M, Phelps D, Duhaney T, Benzies K, Hildebrandt C, Rikhy S, et al. Information and support needs among parents of young children in a region of Canada: a cross-sectional survey. *Public Health Nurs* 2013 May;30(3):193-201. [doi: [10.1111/phn.12002](#)] [Medline: [23586763](#)]
14. Hughes R, Durio HF. Pattern of childcare information seeking by families. *Fam Relat* 1983;32:203-212.
15. Family and Childcare Trust. 2013. National Association of Family Information Services URL: <http://www.daycaretrust.org.uk/nafis> [accessed 2014-03-07] [WebCite Cache ID 6NtAky7Jk]
16. Zubrick SR, Williams AA, Silburn SR, Vimpanic G. Indicators of Social and Family Functioning. Canberra, Australia: Department of Family and Community Services; May 01, 2000:1-68.
17. Karakas F, Lee MD, MacDermid SM. A qualitative investigation into the meaning of family well-being from the perspective of part-time professionals. *Equal Opportunities International* 2004;23:57-77.
18. Chan SS, Viswanath K, Au DW, Ma CM, Lam WW, Fielding R, et al. Hong Kong Chinese community leaders' perspectives on family health, happiness and harmony: a qualitative study. *Health Educ Res* 2011 Aug;26(4):664-674 [FREE Full text] [doi: [10.1093/her/cyr026](#)] [Medline: [21536713](#)]
19. Lam WW, Fielding R, McDowell I, Johnston J, Chan S, Leung GM, et al. Perspectives on family health, happiness and harmony (3H) among Hong Kong Chinese people: a qualitative study. *Health Educ Res* 2012 Oct;27(5):767-779 [FREE Full text] [doi: [10.1093/her/cys087](#)] [Medline: [22907531](#)]
20. Georgas J, Mylonas K, Bafiti T, Poortinga Y, Christakopoulour S, Kagitcibasi C. Functional relationships in nuclear and extended family: a 16-culture study. *International Journal of Psychology* 2001;36:289-300.
21. Wang MP, Wang X, Lam TH, Viswanath K, Chan SS. Health information seeking partially mediated the association between socioeconomic status and self-rated health among Hong Kong Chinese. *PLoS One* 2013;8(12):e82720 [FREE Full text] [doi: [10.1371/journal.pone.0082720](#)] [Medline: [24349347](#)]
22. Census Statistics Department. Hong Kong as an Information Society. 2013. URL: <http://www.digital21.gov.hk/eng/statistics/download/informationssociety2013.pdf> [accessed 2014-09-23] [WebCite Cache ID 6SnO1UZJM]
23. The American Association for Public Opinion Research. Standard Definition Report: Final Dispositions of Case Codes and Outcome Rates for Surveys. Deerfield, IL: AAPOR; 2011:1-61.
24. Viswanath K, Ackerson LK. Race, ethnicity, language, social class, and health communication inequalities: a nationally-representative cross-sectional study. *PLoS One* 2011;6(1):e14550 [FREE Full text] [doi: [10.1371/journal.pone.0014550](#)] [Medline: [21267450](#)]
25. Adler NE, Stewart J. Health disparities across the lifespan: Meaning, methods, and mechanisms. *Ann N Y Acad Sci* 2010 Feb;1186:5-23. [doi: [10.1111/j.1749-6632.2009.05337.x](#)]
26. Clougherty JE, Souza K, Cullen MR. Work and its role in shaping the social gradient in health. *Ann N Y Acad Sci* 2010 Feb;1186:102-124 [FREE Full text] [doi: [10.1111/j.1749-6632.2009.05338.x](#)] [Medline: [20201870](#)]
27. Kawachi I, Adler NE, Dow WH. Money, schooling, and health: Mechanisms and causal evidence. *Ann N Y Acad Sci* 2010 Feb;1186:56-68. [doi: [10.1111/j.1749-6632.2009.05340.x](#)] [Medline: [20201868](#)]
28. Wang MP, Wang X, Lam TH, Viswanath K, Chan SS. The tobacco endgame in Hong Kong: public support for a total ban on tobacco sales. *Tob Control* 2013 Sep 17;In Press. [doi: [10.1136/tobaccocontrol-2013-051092](#)] [Medline: [24046209](#)]
29. Yan YY. Online health information seeking behavior in Hong Kong: an exploratory study. *J Med Syst* 2010 Apr;34(2):147-153. [Medline: [20433053](#)]
30. Fox S, Duggan M. Health online 2013. Washington, DC: Pew Internet & American Life Project; 2013. URL: <http://www.pewinternet.org/2013/01/15/health-online-2013/> [accessed 2014-09-23] [WebCite Cache ID 6SnrRsSv1]
31. Chan YC, Chun PK, Lee TK. Report on the Study of Family Education in Hong Kong. Hong Kong SAR: Department of Applied Social Science; 2012:198.
32. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *JAMA* 2001;285(20):2612-2621. [Medline: [11368735](#)]
33. Steimle BM, Duncan SF. Formative evaluation of a family life education web site. *Fam Relat* 2004;53:367-376.
34. DiMaggio P, Hargittai E. From the Digital Divide to Digital Inequality: Studying Internet Use As Penetration Increases. Princeton, NJ: Princeton University Center for Arts and Cltural Policy Studies; 2001:23.
35. Arcus ME. Family life education: toward the 21st century. *Fam Relat* 1992;41:390-393.

36. Census Statistics Department. 2011 Report on annual earnings and hours survey. Hong Kong SAR: Hong Kong SAR Government; 2012:1-119.
37. FAMILY Project. 2014. Household Survey URL: <http://www.family.org.hk/lang/en-us/household-survey> [accessed 2014-09-23] [[WebCite Cache ID 6SnP9Ed1U](#)]
38. Smilkstein G, Ashworth C, Montano D. Validity and reliability of the family APGAR as a test of family function. *J Fam Pract* 1982 Aug;15(2):303-311. [Medline: [7097168](#)]

Abbreviations

ELIS: everyday life information seeking
FHInTS: Family and Health Information Trends Survey
ICT: information and communication technologies
IRB: Institutional Review Board
NGO: nongovernmental organization
OECD: Organisation for Economic Co-operation and Development
SES: socioeconomic status

Edited by G Eysenbach; submitted 07.03.14; peer-reviewed by A Dubowicz, Q Li; comments to author 17.07.14; revised version received 26.08.14; accepted 13.09.14; published 03.10.14.

Please cite as:

Wang MP, Wang X, Viswanath K, Wan A, Lam TH, Chan SS

Digital Inequalities of Family Life Information Seeking and Family Well-Being Among Chinese Adults in Hong Kong: A Population Survey

J Med Internet Res 2014;16(10):e227

URL: <http://www.jmir.org/2014/10/e227/>

doi: [10.2196/jmir.3386](https://doi.org/10.2196/jmir.3386)

PMID: [25281831](https://pubmed.ncbi.nlm.nih.gov/25281831/)

©Man Ping Wang, Xin Wang, Kasisomayajula Viswanath, Alice Wan, Tai Hing Lam, Sophia S Chan. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 03.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the *Journal of Medical Internet Research*, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Supporting Health Care Professionals to Improve the Processes of Shared Decision Making and Self-Management in a Web-Based Intervention: Randomized Controlled Trial

Barbara Sassen^{1*}, PhD; Gerjo Kok^{2*}, PhD; Jan Schepers^{3*}, PhD; Luc Vanhees^{4*}, PhD

¹Faculty of Health Care, Innovation in Health Care, University of Applied Sciences, Utrecht, Netherlands

²Health Sciences, Work and Social Psychology, Maastricht University, Maastricht, Netherlands

³Health Sciences, Faculty of Psychology and Neurosciences, Maastricht University, Maastricht, Netherlands

⁴Rehabilitation, Rehabilitation Sciences, KU Leuven, Leuven, Belgium

* all authors contributed equally

Corresponding Author:

Barbara Sassen, PhD
Faculty of Health Care
Innovation in Health Care
University of Applied Sciences
Bolognalaan 101
Utrecht, 3584 CJ
Netherlands
Phone: 31 300884815032
Fax: 31 300884815999
Email: Barbara.sassen@hu.nl

Abstract

Background: Research to assess the effect of interventions to improve the processes of shared decision making and self-management directed at health care professionals is limited. Using the protocol of Intervention Mapping, a Web-based intervention directed at health care professionals was developed to complement and optimize health services in patient-centered care.

Objective: The objective of the Web-based intervention was to increase health care professionals' intention and encouraging behavior toward patient self-management, following cardiovascular risk management guidelines.

Methods: A randomized controlled trial was used to assess the effect of a theory-based intervention, using a pre-test and post-test design. The intervention website consisted of a module to help improve professionals' behavior, a module to increase patients' intention and risk-reduction behavior toward cardiovascular risk, and a parallel module with a support system for the health care professionals. Health care professionals (n=69) were recruited online and randomly allocated to the intervention group (n=26) or (waiting list) control group (n=43), and invited their patients to participate. The outcome was improved professional behavior toward health education, and was self-assessed through questionnaires based on the Theory of Planned Behavior. Social-cognitive determinants, intention and behavior were measured pre-intervention and at 1-year follow-up.

Results: The module to improve professionals' behavior was used by 45% (19/42) of the health care professionals in the intervention group. The module to support the health professional in encouraging behavior toward patients was used by 48% (20/42). The module to improve patients' risk-reduction behavior was provided to 44% (24/54) of patients. In 1 of every 5 patients, the guideline for cardiovascular risk management was used. The Web-based intervention was poorly used. In the intervention group, no differences in social-cognitive determinants, intention and behavior were found for health care professionals, compared with the control group. We narrowed the intervention group and no significant differences were found in intention and behavior, except for barriers. Results showed a significant overall difference in barriers between the intervention and the control group ($F_1=4.128$, $P=.02$).

Conclusions: The intervention was used by less than half of the participants and did not improve health care professionals' and patients' cardiovascular risk-reduction behavior. The website was not used intensively because of time and organizational constraints. Professionals in the intervention group experienced higher levels of barriers to encouraging patients, than professionals in the control group. No improvements were detected in the processes of shared decision making and patient self-management.

Although participant education level was relatively high and the intervention was pre-tested, it is possible that the way the information was presented could be the reason for low participation and high dropout. Further research embedded in professionals' regular consultations with patients is required with specific emphasis on the processes of dissemination and implementation of innovations in patient-centered care.

Trial Registration: Netherlands Trial Register Number (NTR): NTR2584; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2584> (Archived by WebCite at <http://www.webcitation.org/6STirC66r>).

(*J Med Internet Res* 2014;16(10):e211) doi:[10.2196/jmir.3170](https://doi.org/10.2196/jmir.3170)

KEYWORDS

Web-based intervention; health professionals; RCT; self-management; barriers

Introduction

In health care, the focus is on optimizing patient self-management. Patients should manage their own health with the support of health care professionals. For targeted and effective self-management, shared decision making is a prerequisite. Shared decision making to improve self-management is more than offering professional support or increasing knowledge about patients' health problem(s). In patient-centered care, patients and health care professionals should cooperate, exchange their own relevant information, and work together optimizing self-management to achieve intended outcomes. It results in better patient outcomes when health care professionals encourage their patients to be involved in decision making. A review showed that professionals tend to misjudge patients' ability to be involved in decision making [1]. Shared decision making is not broadly implemented by health care professionals in clinical practice, and the intention of professionals to engage in or use interventions to facilitate shared decision making is suboptimal [1,2].

To facilitate shared decision making with the objective of optimizing self-management, interventions directed at the health care professional is an option to explore. Intervention Mapping provides a framework to develop systematically planned, theory- and evidence-based interventions [3-6]. Intervention Mapping is used throughout the process of creating an intervention, from diagnosis of the problem to problem solution, and includes collaborating iteratively with priority groups, stakeholders, and experts in the fields of health education and health promotion. Intervention Mapping consists of six planning steps in which each step has a different task and is a prerequisite for the next step. The intervention development process should start by assessing the social-cognitive determinants of the behavior under study. This is followed by choosing and applying methods to change these determinants and behavior. Intervention Mapping places specific emphasis on the transparency of the translation of evidence-based behavior change techniques in intervention components. This is to develop the intervention, explain its rationale, and to facilitate replication [4]. The outcome measures of the intervention should include behavior as well as determinants that influence the behavior [3]. Intervention Mapping has been found to be effective for developing interventions with the objective of changing the behavior of health care professionals and patients, and has led to interventions with effects on patients' behavior [3,7-11].

Information and communication technologies (ICT) in the health care domain (eHealth) can facilitate communication and improve the health of patients and the quality of health care [12,13]. eHealth used as a clinical decision support tool for health care professionals has this potential and can improve clinical practice, though nonusage attrition is a documented problem [12,14,15]. The acceptance of eHealth depends on health care professionals' perception of its usefulness (with an impact on intention), next to the perceived ease of use, and the facilitating and inhibiting conditions, as described in the Technology Acceptance Model [16]. Decision support systems improve clinical practice when these are provided automatically as part of the workflow, in time and at location, and when recommendations are provided [15]. A systematic review showed that computer-based clinical decision support systems can enhance health care professionals' delivered preventive care [17]. In a meta-analysis it was concluded that Web-based instruction for health care professionals had positive effects [18]. But a review of the effectiveness of interventions to promote adoption of ICT showed that there is limited evidence of effective interventions for health care professionals [12]. For patients, Web-based interventions have been effective in boosting health-related changes and improving their health behaviors, though maintenance of the behavior change is often a problem [19]. A meta-analysis of behavioral change outcomes showed that Web-based interventions are effective in achieving changes in knowledge and in the behavior of patients [20]. Web-based interventions that are theory-based and use multiple behavior change techniques, and especially when based on the Theory of Planned Behavior, can change patients' behavior [21,22].

Research to assess the effects of interventions to improve the processes of shared decision making and self-management directed at health care professionals is limited [1]. We hypothesized that in the clinical practice of patient-centered care, shared decision making can optimize self-management using an eHealth-application. In this paper, we report on the results of a randomized controlled trial regarding the implementation of a Web-based intervention directed at health care professionals. The objective of the Web-based intervention was to increase health care professionals' intention and encouraging behavior toward patient self-management, following cardiovascular risk management guidelines [23]. In this paper, we followed the Consolidated Standards of Reporting Trials (CONSORT) criteria [24,25].

Methods

Participants

Participants were health care professionals with at least a bachelor's degree in nursing or physiotherapy and who had regular consultations with patients with cardiovascular risk factors (ie, abdominal obesity, high blood pressure, low high-density lipoprotein cholesterol, elevated triglycerides, and elevated blood glucose levels) and low levels of physical activity [26-29]. The health care professionals were former students of the Faculty of Health Care, University of Applied Sciences (Netherlands) and were invited to participate via a personalized email. All participants were offered a 3-hour training session to work with the website and 12 participants attended the training. Participants could choose not to attend the training and instead use the online instruction tool. Much effort was put in motivating the participants in the intervention group to use and keep using the website. For follow-up contact, we used emails and telephone contact.

Study Design

A randomized controlled trial was performed using a pre-test (T1) and post-test (T2) design, to determine the effectiveness of a clinical decision support system used to optimize shared decision making and the self-management of patients. Health care professionals were the unit of randomization and were randomly allocated to the intervention versus the waiting list control group. Health care professionals invited their patients. The recommendation was to use the intervention for every patient that fit the intervention. Patients were informed about the study and gave informed consent. The study sample size was based on the outcome of improved professionals' behavior toward patient-centered care. Power analysis estimated how many respondents were needed for the study to find a significant difference in health care professionals' behavior. This analysis (power 0.80; $\alpha=.05$, two-tailed) revealed that 62 professionals in each condition were needed. Randomization was based on a random number sequence, using a computer randomized number generator. The total group of 278 professionals was randomized and drop by drop assigned to the intervention versus the control group; 81 professionals were willing to participate (Figure 1). The professionals in the waiting list control group did not receive further intervention, until after T2. The professionals in the control group did not receive a log-in code, and did not invite or enlist patients. A total of 12 professionals were "lost", 9 for not using their log-in code and 3 of them changed their email accounts. Data collection via the website took place at T1 and T2 after 12 months. Initially the follow-up period was 6 months; however, due to low participation rates, we extended the follow-up period. Professionals used the website from January 2011 till June 2012.

Outcomes were self-assessed through a questionnaire based on the Theory of Planned Behavior [30,31]. The questionnaire was part of the website and used at T1 and T2. The content was derived from a literature review and in-depth interviews with health care professionals on how they encourage patients. There were eight elicitation interviews held, four with health care professionals with a background in physiotherapy and four with

a background in nursing. Four of those professionals were observed in their professional activities for a regular working day. For measuring social-cognitive determinants, no valid questionnaires are available, but only valid procedures. The construction of the questionnaire is according to the Theory of Planned Behavior, specific to the definition of the behavior and the specification of the research population [4,31]. The questionnaire was piloted and as a result no revisions were made. Answers ranged on a 7-point scale (1="definitely not" to 7="most definitely").

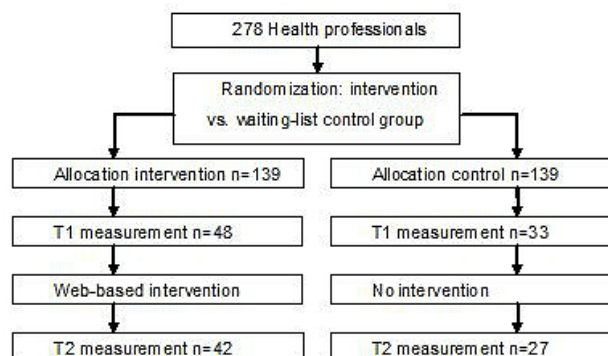
We assessed "behavior" with two items: "Do you encourage cardiovascular patients to increase their physical activity?", and "How often do you encourage cardiovascular patients to become physically active?" (Cronbach $\alpha=.64$). "Intention" was indexed with three questions: "Do you intend to encourage cardiovascular patients to become physically active tomorrow and the day after tomorrow?", "Do you expect to encourage cardiovascular patients to become physically active tomorrow and the day after tomorrow?", and "Of the first 10 cardiovascular patients you see, how many do you intend to encourage to become physically active?" (Cronbach $\alpha=.82$). "Attitude" was assessed by: "In my view, encouraging cardiovascular patients to become physically active is very good - very bad" and "Encouraging cardiovascular patients is very useful - very useless". Then we asked, "Is it useful to: assess patients' motivation, assess the pros and cons of physical activity, teach patients how to resist social pressure, teach patients specific skills pertaining to physical activity, teach patients how to handle barriers in regard to physical activity, formulate physical activity goals together with patients, teach patients how to handle relapses, and help patients understand the relationship between the specific health problem and physical inactivity?". These eight items ("Is it useful to...") were averaged and that score was averaged with the first two item scores to represent attitude (Cronbach $\alpha=.63$).

"Perceived behavioral control" (PBC) was assessed by: "Do you think that you have the skills and knowledge to encourage cardiovascular patients to become physically active?", "Do you think you can rely on your skills and knowledge to encourage cardiovascular patients to become physically active?", and third we asked, "Encouraging every cardiovascular patient to become physically active is very difficult (1) - very easy (7)". PBC was further assessed by eight items that paralleled the eight items used for attitudes: "It is very difficult (1) - very easy (7) to: assess patients' motivation, assess the pros and cons of physical activity, teach patients how to resist social pressure, teach patients specific skills pertaining to physical activity, teach patients how to handle barriers in regard to physical activity, formulate physical activity goals together with patients, teach patients how to handle relapses, and help patients understand the relationship between the specific health problem and physical inactivity?". Once again, this scale score was calculated and combined with the previous three items as a measure of PBC (Cronbach $\alpha=.68$). "Subjective norm" was measured by four items: "Most colleagues who are important to me think I should encourage cardiovascular patients to become physically active", "Most colleagues value that I encourage cardiovascular patients to become physically active", "Patients value that I

encourage them to become physically active”, and “The organization I work for values that I encourage cardiovascular patients to become physically active” (Cronbach alpha=.73). “Moral norm” was assessed by three questions: “Encouraging patients to engage in physical activity is...my professional duty, ...a moral obligation, and ...an obvious part of my job” (Cronbach alpha=.76). A determinant score was calculated for

every social-cognitive determinant. “Habit” was measured by two questions: “Encouraging patients to be physically active is something I do without thinking, and ...something I do automatically” (Cronbach alpha=.75). “Barriers” were indexed by two questions that focused on encouraging patients “even when one is busy” and “even when one’s organization makes it difficult to encourage patients” (Cronbach alpha=.69).

Figure 1. Intervention flow chart.



Intervention

Participants had access to the website [32], which offered several modules. The development and content of the intervention is described in detail elsewhere [7,30]. Modules and a forum were directed at the health care professional to increase professionals’ awareness of their thoughts, and learn skills and strategies to support patients in their own self-management, this to improve their intention and behavior toward patient-centered health education. The first module enclosed a set of seven screens to help the professional to improve his or her professional behavior (Figure 2). The screens contained self-complete forms and were designed and pre-tested to educate the health care professional, with a personal feedback system in a “coaching spider chart”. The screens started with “risk” communication to support thinking about encouraging patients. This was followed by listing the pros and cons of encouraging patients in the short- and long-term. Hereafter, the health professional was encouraged to seek support and look at the sub-skills needed to be an encouraging health professional. Next, there was a screen for planning the encouraging behavior change, making a plan, and putting the behavior change into practice. The identification of high-risk situations and the practice of coping responses were encouraged. To enhance effective patient-centered health education, the website also included a second module with a support system for the health professional, parallel to the module for encouraging the patient (Figure 3). The third module consisted of a maximum of seven consultations to encourage the patient with cardiovascular risk factors, easily adaptable to

the needs and individual characteristics of the patients (Figure 4). This module started with risk perception to encourage the patient to think about individual cardiovascular risk and personal vulnerability, followed by encouraging the patient to describe what the personal pros and cons are (not) to becoming physically active in the short- and long-term. With the support of the professional, the patient was encouraged to recognize social pressure, seek social support, and practice sub-skills. The patient was supported in planning the behavior and putting it into practice, detecting high-risk situations, and practicing coping responses. The third module started with the assessment following cardiovascular risk management guidelines. The screens contained the patient’s profile with a feedback system on the progression in behavior change in a spider chart, physical activity levels in bar graphs, and cardiovascular risk factors in a pie chart. The website helped the patient to look back at the plans made in conjunction with the health care professional. The website provided a fourth module with specific information on physical activity devices, planning physical activity, and cardiovascular risk factors. The website also included a link to a forum directed at health care professionals to share experiences with other professionals in the intervention group. The intervention was extensively tested, but not piloted. The website underwent no changes during implementation. Institutional affiliations were displayed on the website. The study was approved by the Research Ethics Board at Maastricht University and was registered in the Dutch Trial Register (Trial ID: ECP-92, NTR2584).

Figure 2. Intervention screenshot.

Gezondheidsbevordering.eu
Brengt je in beweging

Home Patiënt Motivering Professional Forum Activiteit & Fitheid

Profiel professional

Coachingsspin

CB, AV, SO, EE, IV, GV, RP

■ Initie
■ Huidig
■ Maximum

Gezondheidscoach - Attitudeverandering

Een patiënt motiveren tot ander, gezonder gedrag heeft zo zijn voor- en nadelen. Wat zijn voor jou persoonlijk de voordelen om een patiënt te motiveren? In je profiel professional kun je zien hoe het staat met je attitude (AV). Wat zijn voor jou persoonlijk de nadelen om een patiënt te motiveren?

	KORT TERMIJN	LANGE TERMIJN
Voordelen	1: 2: 3: 4:	1: 2: 3: 4:
Nadelen	1: 2: 3: 4:	1: 2: 3: 4:

Wil je op de onderstaande schaal (0 t/m 10) de mate aangeven waarin je door deze website te gebruiken een betere gezondheidscoach wilt worden.

Niet willen veranderen 2 3 4 5 6 7 8 9 10 Ik wil een betere gezondheidscoach worden

Opslaan & vorige Opslaan & volgende

HOGESCHOOL UTRECHT Universiteit Maastricht NEDERLANDSE UNIVERSITEIT LEUVEN

© Copyright 2010 Gezondheidsbevordering.eu — Alle rechten voorbehouden — Disclaimer

Figure 3. Intervention screenshot.

Gezondheidsbevordering.eu
Brengt je in beweging

Home Patiënt Motivering Professional Forum Activiteit & Fitheid

Motivering bij stap 1: Risicoperceptie

Selecteer de tekst die je belangrijk vindt door deze aan te klikken.

Doelen

Doel 1

Je doet je intake bij de patiënt zoals je dat altijd doet en combineert dit met vragen vanuit de Richtlijn Cardiovasculair Risicomanagement.

Doel 2

Je werkt eraan dat de patiënt het verband begrijpt tussen het gezondheidsprobleem (risicofactoren) en de gewenste gedragsverandering (meer bewegen). De patiënt maakt een persoonlijke risico-inschatting, accepteert dit en onthoudt het.

Achtergronden

Als een patiënt voldoende kennis heeft over de relatie tussen het gezondheidsprobleem en fysieke activiteit en fitheid én een goede inschatting maakt van het persoonlijke risico (zgn. risk-perception) is de kans van slagen groter. Het gaat om de inschatting die een persoon maakt van het risico dat hij of zij loopt op het gezondheidsprobleem. In het algemeen onderschat de patiënt het eigen, individuele risico en denkt de patiënt dat anderen het risico lopen op het gezondheidsprobleem. Kennisvermeerdering leidt in de regel nooit tot gedragsverandering en is op zichzelf onvoldoende om aan te zetten tot (meer) bewegen.

Voorbeelden

Doel 1

"In dit gesprek gaat het erom dat ik u wat beter leer kennen en dat u weet wat u van mij kunt verwachten. We starten met een algemene intake, hierna zou ik wat dieper in willen gaan op uw gezondheidsprobleem". Stel je vragen.

Doel 2

"U bent hier gekomen (met verwijzing, etc). Kunt u mij vertellen wat u al weet over (gezondheidsprobleem)?" (Patiënt vertelt)
Vervolg: "Het klopt wat u zegt (over het gezondheidsprobleem) namelijk dat (herhaal hetgeen de patiënt vertelde over het gezondheidsprobleem). Denkt u dat u zelf een grote kans loopt op (dit gezondheidsprobleem)?"
"Ik begrijp dat het niet prettig is om te horen, maar er is reden tot bezorgdheid als (u uw leefstijl) niet zou aanpassen/meer gaat bewegen. Dit hoeft geen onoverkomelijk probleem te zijn, ik kan u helpen uw (leefstijl) te veranderen."

Algemeen

Bij Notities kun je (voor risicoperceptie) notities maken. Leefstijladviezen worden in de regel nu nog niet gegeven, maar volg je eigen routine.

Stel steeds open vragen om de betrokkenheid van de patiënt te vergroten, geeft patiënt het gevoel dat hij een waardevolle gesprekspartner is.

HOGESCHOOL UTRECHT Universiteit Maastricht RADBOD UNIVERSITEIT LEUVEN

© Copyright 2010 Gezondheidsbevordering.eu — Alle rechten voorbehouden — Disclaimer

Figure 4. Intervention screenshot.

The screenshot displays the 'Gezondheidsbevordering.eu' website interface. The header includes the logo and the tagline 'Brengt je in beweging'. A navigation bar contains links for Home, Patient, Motivering, Professional, Forum, and Activiteit & Fitheid.

The main content area is divided into two columns. The left column shows the 'Profiel patiënt' for James Window, aged 58. It includes a 'Risicofactoren' pie chart (Gezond, Middelen, HDL, Glucose, Triglyceride, Bloeddruk) and a 'Fysieke activiteit & Fitheid' bar chart comparing 'Intake' and 'Huidig' levels across various activities. A 'Gezondheidsspin' radar chart shows performance across different health domains. The bottom of the left column indicates that no movement advice has been selected.

The right column is titled 'Stap 1: Risicoperceptie - Waarom bewegen'. It features a 'Toon suggesties' button and a list of goals, including understanding cardiovascular risk factors and increasing physical activity. Below this are two text input fields for 'Notities' and 'Activiteit & Fitheid'. A 'Bewegingsadvies' section lists 11 options, such as 'Actieve leefstijl in het dagelijks leven' and '15 minuten matig intensief lichaamelijk actief zijn, 3 dagen per week'. At the bottom of the right column are 'Opslaan & vorige' and 'Opslaan & volgende' buttons.

The footer contains logos for Hogeschool Utrecht, Universiteit Maastricht, and Middelburgse Hogeschool Leuven, along with a copyright notice for 2010.

Statistical Analysis

Descriptive statistics were calculated and chi-square analyses were used to characterize the study groups at baseline and to determine the use of the website. We used paired *t* tests to evaluate the differences at T1 and T2, for the intervention group and the control group. Subsequently, we applied the method of General Linear Modeling with repeated measures to explore the overall change between the intervention and control group. Because we noticed from the process evaluation data that 16 health care professionals in the intervention group did not start using the website, we applied the method of General Linear

Modeling with repeated measures again; for this analysis, these 16 professionals were transferred to the waiting list control group. IBM SPSS version 20.0 was applied and a *P* value of $\leq .05$ was considered as statistically significant.

Results

Table 1 shows the baseline demographics of the study participants in the intervention and waiting list control group. A total of 69 health care professionals were identified, with 42 professionals participating in the intervention group: 69% (29/42) female, mean age 38.6 years (SD 11.3); 79% bachelor's

degree, 21% higher education; with mean 9.76 years (SD 8.5) of professional experience; 44% worked as a soloist or with 1 or 2 colleagues, and 56% worked together with at least 3 others). Professionals in the intervention group stated that 59% of their consultation time was devoted to health education. No significant differences between the intervention group and the control group (n=27) were found at baseline (T1).

Figure 5 provides a representation of the use of different modules on the website: a module to improve professional behavior, a module to support the health professional, and a module to improve patients' intention and risk-reduction behavior. The module to improve the professionals' behavior to optimize processes of shared decision making and self-management, was used by 45% of the professionals (19/42). Of the professionals in the intervention group, 19% (8/42) used only one of seven screens; 7% (3/42) used all seven screens. The screen to support thinking about encouraging patients was used by 38% (16/42) of the professionals. In 19% (8/42) of the cases, pros and cons of encouraging patients in the short- and long-term were listed, 14% (6/42) used the screen to seek support, and 17% (7/42) looked at the sub-skills needed. In addition, the screen for planning the encouraging behavior was used by 17% (7/42) of the professionals, for putting the behavior change into practice by 10% (4/42), and for maintaining the encouraging behavior, 17% (7/42).

The module with background information on how to coach the patient with the aim of supporting the health professional in his or her encouraging behavior toward patients was used by 48% (20/42) of the professionals; 17% (7/42) of the professionals used all seven screens; 45% (19/42) used the screen how to encourage a patient to think about his/her personal risk; 33% (14/42) used the screen how to list pros and cons with a patient; 26% (11/42) used the screen how to seek support; and 21% (9/42) how to practice the sub-skills needed. The last screens in this module (planning the encouraging behavior, putting the behavior change into practice, and maintaining the behavior) were used by 19% (8/42) of the health care professionals. The forum directed at improving social support was used by 4 health care professionals.

Health care professionals invited 54 patients in the intervention: 56% (30/54) male, mean age 50.9 years (SD 11.8), with differing educational degrees. Health care professionals assessed cardiovascular risk and 19% (10/54) of the patients had two or more cardiovascular risk factors, and/or a heart disease and/or diabetes; 13% (7/54) were physically active for at least 30 minutes, for 5 days per week. In 82% (44/54) of the patients, the guidelines to assess cardiovascular risk were not used by the health professional. The module to improve patients' intention and risk-reduction behavior, with the purpose of increasing the processes of shared decision making and self-management, was provided to 44% (24/54) of the patients by a health professional (Figure 2). Of the professionals that

used this module together with their patients, 43% (23/54) provided their patient(s) with risk-perception. A total of 39% (21/54) provided attitudinal change and outcome expectations, 30% (16/54) resistance to social pressure and seek support, and 28% (15/54) provided encouraging sub-skill enactment. For 24% (13/54) of the patients, the professional provided a planning of the behavior change; 19% (10/54) provided putting the behavior change into practice; and 15% (8/54) provided maintaining the behavior change.

Table 2 shows the baseline score at T1 and T2, for the intervention group and the control group. At measurement T1, the professionals in the intervention group had positive intentions (mean 6.25, SD 1.00) and positive attitudes (mean 6.30, SD 0.44) to encouraging patients with cardiovascular risk factors, but, in comparison, the mean scores on the self-reported behavior (mean 4.54, SD 1.02), perceived behavioral control (mean 4.65, SD 0.79), and subjective norms (mean 5.48, SD 0.55) were modest. At T1, the mean intention score was 6.25 (SD 1.00) and the behavior score was 4.54 (SD 1.02); at T2, the mean intention score was 6.06 (SD 1.11) and behavior score was 4.63 (SD 0.85). In the intervention group, no differences in social-cognitive determinants, intention and behavior were found when we compared T1 and T2, except for a difference in perceived behavioral control ($t_{26}=-2.954$, $P=.00$, effect size=0.50). In the control group, also no differences were detected between the measurement at T1 compared with T2; but, contrary to expectations, we detected an increase in perceived behavioral control ($t_{19}=-2.651$, $P=.02$, effect size=0.54). When we compared the intervention group with the professionals in the control group, no significant differences between the intervention and control group were found in social-cognitive determinants, intention and behavior; the detected difference in the intervention and control group in perceived behavioral control turned out to be overall non-significant when we compared the intervention with the control group.

When we narrowed the intervention group (n=26) by transferring the professionals who did not use the website to the waiting list control group (n=43), no significant differences between the intervention and control group were found in social-cognitive determinants, intention and behavior, except for perceived behavior control and barriers (Table 3). There was a difference in perceived behavioral control in the intervention group ($t_{19}=-2.485$, $P=.02$, effect size=-0.30), and also in the control group ($t_{18}=-3.105$, $P=.00$, effect size=-0.23). This detected difference in perceived behavioral control turned out to be overall non-significant. Results showed a significant overall difference in barriers between the intervention and the control group ($F_1=4.128$, $P=.02$). Professionals in the intervention group experienced higher levels of barriers to encouraging patients, than professionals in the control group.

Table 1. Baseline demographics of study participants (N=69).

	Intervention group (n=42) n (%)	Control group (n=27) n (%)	P value
Gender, female	29 (69%)	21 (78%)	.428
Age, years, mean (SD)	38.6 (11.3)	39.7 (8.4)	.062
Education, bachelor's degree	26 (79%)	13 (68%)	.406
Education, degree above bachelor's	7 (21%)	6 (32%)	N/A
Professional experience, years, mean (SD)	9.76 (8.5)	9.58 (9.1)	.910
Working as a soloist, or with 1 or 2 colleagues	16 (44%)	5 (21%)	.060
Working with 3 or more colleagues	20 (56%)	19 (79%)	N/A
Consultation time devoted to health education	36 (59%)	23 (54%)	.508

Table 2. Paired differences between intervention group and control group, measured at T1 and T2.^a

		Intervention group (n=42) mean (SD)	Control group (n=27) mean (SD)	F test
Behavior				
	T1	4.54 (1.02)	4.83 (0.69)	P=.68
	T2	4.63 (0.85)	4.79 (0.82)	
Intention				
	T1	6.25 (1.00)	5.87 (1.15)	P=.12
	T2	6.06 (1.11)	6.02 (0.91)	
Attitude				
	T1	6.30 (0.44)	6.23 (0.69)	P=.64
	T2	6.30 (0.56)	6.31 (0.68)	
Perceived behavioral control^b				
	T1	4.65 (0.79)	4.90 (0.87)	P=.98
	T2	5.04 (0.73)	5.28 (0.80)	
Subjective norms				
	T1	5.48 (0.55)	5.58 (0.93)	P=.64
	T2	5.57 (0.63)	5.74 (0.76)	
Moral norms				
	T1	6.04 (0.63)	6.20 (0.59)	P=.75
	T2	6.19 (0.70)	6.30 (0.55)	
Barriers				
	T1	3.11 (1.17)	2.78 (1.01)	P=.46
	T2	3.18 (1.12)	2.63 (0.96)	

^asocial-cognitive variables range 1-7.

^bIntervention group: Cohen's $d=-1.15$, effect size=0.50; Control group: Cohen's $d=-1.28$, effect size= 0.54.

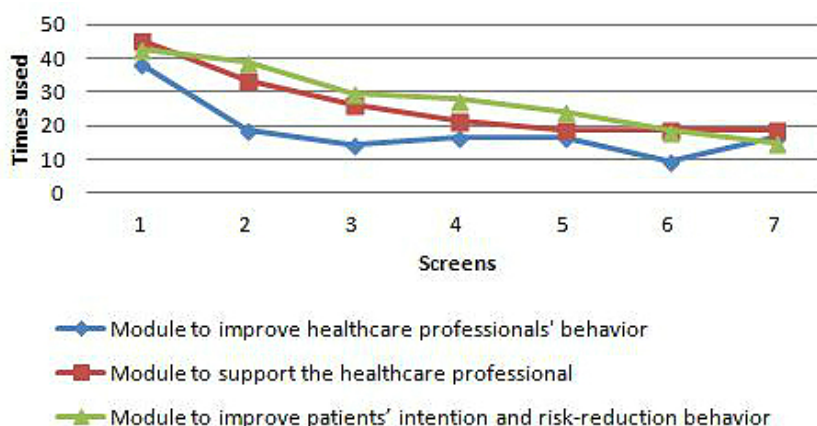
Table 3. Differences between narrowed intervention group and control group, measured at T1 and T2.^a

		Intervention group (n=26) mean (SD)	Control group (n=43) mean (SD)	F test
Behavior	T1	4.38 (1.04)	4.91 (0.67)	<i>P</i> =.76
	T2	4.45 (0.81)	4.89 (0.81)	
Intention	T1	6.10 (1.06)	6.08 (1.10)	<i>P</i> =.33
	T2	5.93 (1.21)	6.13 (0.85)	
Attitude	T1	6.28 (0.47)	6.26 (0.63)	<i>P</i> =.86
	T2	6.29 (0.56)	6.31 (0.64)	
Perceived behavioral control ^b	T1	4.63 (0.82)	4.85 (0.83)	<i>P</i> =.96
	T2	5.01 (0.76)	5.23 (0.77)	
Subjective norms	T1	5.46 (0.59)	5.57 (0.84)	<i>P</i> =.91
	T2	5.60 (0.63)	5.68 (0.73)	
Moral norms	T1	5.88 (0.60)	6.28 (0.58)	<i>P</i> =.41
	T2	6.08 (0.73)	6.35 (0.53)	
Barriers	T1	3.09 (1.11)	2.78 (1.11)	<i>F</i> ₁ =4.128 <i>P</i> =.05
	T2	3.40 (1.09)	2.59 (0.94)	

^asocial-cognitive variables range 1-7.

^bIntervention group: Cohen's *d*=-0.63, effect size=-0.30; Control group: Cohen's *d*=-0.47, effect size=-0.23.

Figure 5. Use of the intervention website modules.



Discussion

Principal Findings

In this paper, we report on the results of a randomized controlled trial testing the effectiveness of a Web-based intervention in

the clinical practice of patient-centered care. The intervention was developed to optimize processes of shared decision making and self-management, following the protocol of Intervention Mapping. The objective was to increase health care professionals' intention and behavior toward encouraging patient self-management.

Results indicated no intervention effect on the outcome measure of our study: the encouraging behavior of health care professionals. Results indicated no overall differences for social-cognitive determinants, intention and behavior, when the intervention group was compared with the control group. We narrowed the intervention group and took a closer look at the health care professionals that used the Web-based intervention. Results showed that these professionals experienced higher levels of barriers, meaning that time and organizational constraints withheld them and obstructed the planned behavior to encourage patients, compared with the professionals in the control group. Next to the overall results of the intervention, we took a closer look at possible effects in the (initial) intervention group. Results indicated a medium-size effect for perceived behavioral control, with no effect for the other social-cognitive determinants, intention and behavior. Professionals in the intervention group increased their perceived behavioral control and reported that they had more control over their skills necessary to encourage patients. The same effect was seen in the control group, which means that there was no overall effect when we compared the intervention group with the control group.

Our study showed that health care professionals had high intentions and planned their encouraging behavior. It also showed that health care professionals had positive attitudes and described more pros than cons toward encouraging patients. Further, the study showed a positive moral norm to be an encouraging professional. But scores on behavior were modest in comparison, and though health care professionals did plan the encouraging behavior, they did not practice the encouraging behavior. Also scores on subjective norm (meaning that colleagues, patients, and the organization value their encouragement) and scores on perceived behavior control as the skills needed, were modest. Attendance on the Web-based intervention and use of the website was sub-optimal. Less than half of the health care professionals used the module to change their professional behavior, and/or used the module to get support in their encouraging behavior toward patients, and/or used the module to improve patients' intention and risk-reduction behavior. The module to improve patients' intention and risk-reduction behavior was used most, followed by the module to support the health professional. The module to change professional behavior was the least used. Only in 1 of every 5 patients was the guideline following cardiovascular risk management used. We hypothesized that in the clinical practice of patient-centered care, shared decision making can optimize self-management using an eHealth-application, but we were not able to detect improvements in the processes of shared decision making and self-management of the patient.

Systematic reviews showed a clear relationship between the intentions of health care professionals and their subsequent behavior; these were found to be appropriate to predict their behavior, and can be used to improve behavior change interventions targeting health care professionals [33,34]. Although the intention to employ interventions to facilitate shared decision making is often suboptimal, the health care professionals in our study showed high intention scores [1,2]. If intention scores are already high at baseline, it is difficult to

change the behavior scores in a positive direction. Research showed that medium-to-large changes in intention scores are needed to show small-to-medium changes in behavior [35]. That improvement in intention was not sufficient to change the behavior, is often a problem and can be related to the fact that the effectiveness of interventions is reduced with increasing levels of standard care [36,37]. We detected an increase in perceived behavioral control and other research showed that health care professionals' perceived behavior control is an important determinant of behavior to improve shared decision making [1]. The improvement in perceived behavioral control is important, because this can lead to progress in (planning) the encouraging behavior. But it is unclear if we can attribute the change in perceived behavioral control as an intervention effect, because we also detected an increase in the control group. In our study, health care professionals experienced barriers that hinder them in their encouraging behavior, meaning that time and organizational constraints obstruct them in improving the processes of shared decision making and self-management in patient-centered care. Another intervention directed at optimizing shared decision making also showed that lack of time was a barrier although intentions were high [1,38]. Other research showed that barriers for using an intervention are also that health care professionals do not use the applications, poor usability or integration into professionals' workflow, non-acceptance of recommendations, and also the intervention's inapplicability due to patient characteristics and the clinical situation [1,17,38]. In our study, changing the health behavior in line with evidence-based recommendations as described in the guideline for cardiovascular risk management proved difficult, but was similar with other studies where only small or no effects were found [39].

The application of evidence-based behavior change techniques used in our intervention should offer insight regarding how an intervention may change intention and behavior. When intention and perceived behavioral control are targeted in an intervention, clinician behavior can be improved [35]. Methods used in our Web-based intervention were action planning and coping planning; however, better results on intention and the (maintenance) of the behavior change were not reached [36]. Professionals' perception of perceived behavioral control is an important determinant of behavior to improve shared decision making [1]. The use of the method-guided practice with feedback probably did lead to increased perceived behavioral control and better skills. But, use of the method decisional balance to encourage listing of pros and cons of changing the behavior did not lead to better results on attitudes. The use of the method resistance to social pressure and mobilizing others for social support, showed a slight but non-significant increase in subjective norm. We showed that professionals did not use or discontinued using the Web-based intervention. Though health care professionals stated that they spend a lot of time on health education (58.4% in the intervention group and 54.4% in the control group), and the pre-test before implementation of the intervention showed that a more evidence-based and systematic approach could be done in the same time, health care professionals lost interest in the intervention and stopped using the application. This nonuse of the intervention or nonusage attrition is a documented problem in the search for intervention

effects [14]. Eysenbach [14] called this the methodological challenge in the evaluation of eHealth applications. A clinical decision support system can improve health care professionals' performance when users are automatically prompted to use the website, but in our study the professionals themselves had to initiate use of the system [17]. A factor that may have influenced performance and attrition may be a lack of (immediate) advantage of working with the website for the health care professionals, or even encountering obstructions when working with the website. Another factor that may have influenced performance and attrition may be the compatibility with usual care and with workload. Also, the complexity of the intervention with (too) many modules may have influenced performance and attrition. The Web-based intervention was carefully developed following the Intervention Mapping protocol [7]. Health care professionals from the target group were involved in the development of the intervention, and they also pre-tested the application, but it may be that (too) much attention was directed at "does the website work as intended".

Limitations

In a review by Légare, it was concluded that sufficient enrollment of health care professionals and patients is often a problem and needs attention in research designs [1]. Our estimated sample size was not achieved, although recruitment and follow-up period was extended, and strong attempts were made to encourage professionals. Changing intentions and the encouraging behavior of health care professionals proved difficult with many inhibiting factors. An explanation is probably a ceiling effect for intention and attitude, so little progress can be expected in the intervention group. This combined with the fact that our measures were based on self-reported intention and behavior can have caused recall bias. Also, a possible explanation is a selection of professionals with special interest in health education, participating in our study. Although randomization ensured that participants were evenly distributed, we noticed that a few professionals were trained for another intervention but were not selected and they choose to attend our intervention. A limitation that might also have influenced outcomes is that the website had to be used during regular consultation with the patient. Though professionals

stated that they spend much of their consultation time on health education, it was not just a matter of focus on self-management during consultation time. Another limitation that might influence the outcome is that the professionals needed more training on how to work with the modules of the Web-based intervention.

An online intervention can support health care professionals but training should be an important part of implementation. A total of 12 health care professionals attended a demonstration meeting, including professionals in the waiting list control group. It may be that training on the job can improve the use of the Web-based intervention. Training may increase professionals' perception of perceived behavioral control, because professionals need to learn to use the specific clinical decision support tool [1,12,40]. Training that uses practice exercises, repetition, and feedback leads to improved learning outcomes for health care professionals [18,39]. Training may improve the process of shared decision making and self-management, as may the implementation of patient-mediated interventions such as decision aids [38]. Other important facilitators for dissemination and implementation of innovations are increasing health care professionals' motivation, and showing the intervention's innovative impact on the clinical process and on patient outcomes [1,41,42].

Conclusions

The intervention was used by less than half of the participants and did not improve health care professionals' and patients' cardiovascular risk-reduction behavior. Health care professionals did not use the website intensively because of time and organizational constraints. Professionals in the intervention group experienced higher levels of barriers to encouraging patients, than professionals in the control group. We were not able to detect improvements in the processes of shared decision making and patient self-management. Although participant education level was relatively high and the intervention was pre-tested, it is possible that the way the information was presented could be the reason for low participation and high dropout. Further research embedded in professionals' regular consultations with patients is required with specific emphasis on the processes of dissemination and implementation of innovations in patient-centered care.

Acknowledgments

The authors would like to thank Ir Lex Verheesen for his contribution in the development of the intervention website.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [25].

[PDF File (Adobe PDF File), 1010KB - [jmir_v16i10e211_app1.pdf](#)]

References

1. Légaré F, Ratté S, Stacey D, Kryworuchko J, Gravel K, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. *Cochrane Database Syst Rev* 2010(5):CD006732. [doi: [10.1002/14651858.CD006732.pub2](https://doi.org/10.1002/14651858.CD006732.pub2)] [Medline: [20464744](https://pubmed.ncbi.nlm.nih.gov/20464744/)]
2. Elwyn G, Edwards A, Britten N. "Doing prescribing": how might clinicians work differently for better, safer care. *Qual Saf Health Care* 2003 Dec;12 Suppl 1:i33-i36 [FREE Full text] [Medline: [14645747](https://pubmed.ncbi.nlm.nih.gov/14645747/)]
3. Parcel GS, Kok G, Gottlieb NH. *Planning Health Promotion Programs: An Intervention Mapping Approach*. San Francisco: Jossey-Bass; 2011.
4. Kok G, Schaalma H, Ruiters 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](https://doi.org/10.1177/1359105304038379)] [Medline: [14683571](https://pubmed.ncbi.nlm.nih.gov/14683571/)]
5. Schaalma H, Kok G. Decoding health education interventions: the times are a-changin'. *Psychol Health* 2009 Jan;24(1):5-9. [doi: [10.1080/08870440903126348](https://doi.org/10.1080/08870440903126348)] [Medline: [20186636](https://pubmed.ncbi.nlm.nih.gov/20186636/)]
6. Kok G, Mesters I. Getting inside the black box of health promotion programmes using intervention Mapping. *Chronic Illn* 2011 Sep;7(3):176-180. [doi: [10.1177/1742395311403013](https://doi.org/10.1177/1742395311403013)] [Medline: [21900338](https://pubmed.ncbi.nlm.nih.gov/21900338/)]
7. 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](https://doi.org/10.2196/resprot.1804)] [Medline: [23612470](https://pubmed.ncbi.nlm.nih.gov/23612470/)]
8. van Stralen MM, Kok G, de Vries H, Mudde AN, Bolman C, Lechner L. The Active plus protocol: systematic development of two theory- and evidence-based tailored physical activity interventions for the over-fifties. *BMC Public Health* 2008;8:399 [FREE Full text] [doi: [10.1186/1471-2458-8-399](https://doi.org/10.1186/1471-2458-8-399)] [Medline: [19055806](https://pubmed.ncbi.nlm.nih.gov/19055806/)]
9. Wolfers ME, van den Hoek C, Brug J, de Zwart O. Using Intervention Mapping to develop a programme to prevent sexually transmittable infections, including HIV, among heterosexual migrant men. *BMC Public Health* 2007;7:141 [FREE Full text] [doi: [10.1186/1471-2458-7-141](https://doi.org/10.1186/1471-2458-7-141)] [Medline: [17615052](https://pubmed.ncbi.nlm.nih.gov/17615052/)]
10. Heinen MM, Bartholomew LK, Wensing M, van de Kerkhof P, van Achterberg T. Supporting adherence and healthy lifestyles in leg ulcer patients: systematic development of the Lively Legs program for dermatology outpatient clinics. *Patient Educ Couns* 2006 May;61(2):279-291. [doi: [10.1016/j.pec.2005.04.007](https://doi.org/10.1016/j.pec.2005.04.007)] [Medline: [15964733](https://pubmed.ncbi.nlm.nih.gov/15964733/)]
11. Alewijnse D, Mesters IE, Metsemakers JF, van den Borne BH. Program development for promoting adherence during and after exercise therapy for urinary incontinence. *Patient Educ Couns* 2002 Nov;48(2):147-160. [Medline: [12401418](https://pubmed.ncbi.nlm.nih.gov/12401418/)]
12. Gagnon MP, Légaré F, Labrecque M, Frémont P, Pluye P, Gagnon J, et al. Interventions for promoting information and communication technologies adoption in healthcare professionals. *Cochrane Database Syst Rev* 2009(1):CD006093 [FREE Full text] [doi: [10.1002/14651858.CD006093.pub2](https://doi.org/10.1002/14651858.CD006093.pub2)] [Medline: [19160265](https://pubmed.ncbi.nlm.nih.gov/19160265/)]
13. Eysenbach G. What is e-health? *J Med Internet Res* 2001;3(2):E20 [FREE Full text] [doi: [10.2196/jmir.3.2.e20](https://doi.org/10.2196/jmir.3.2.e20)] [Medline: [11720962](https://pubmed.ncbi.nlm.nih.gov/11720962/)]
14. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
15. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005 Apr 2;330(7494):765 [FREE Full text] [doi: [10.1136/bmj.38398.500764.8F](https://doi.org/10.1136/bmj.38398.500764.8F)] [Medline: [15767266](https://pubmed.ncbi.nlm.nih.gov/15767266/)]
16. Li J, Talaie-Khoei A, Seale H, Ray P, Macintyre CR. Health care provider adoption of eHealth: systematic literature review. *Interact J Med Res* 2013;2(1):e7 [FREE Full text] [doi: [10.2196/ijmr.2468](https://doi.org/10.2196/ijmr.2468)] [Medline: [23608679](https://pubmed.ncbi.nlm.nih.gov/23608679/)]
17. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005 Mar 9;293(10):1223-1238. [doi: [10.1001/jama.293.10.1223](https://doi.org/10.1001/jama.293.10.1223)] [Medline: [15755945](https://pubmed.ncbi.nlm.nih.gov/15755945/)]
18. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. *JAMA* 2008 Sep 10;300(10):1181-1196. [doi: [10.1001/jama.300.10.1181](https://doi.org/10.1001/jama.300.10.1181)] [Medline: [18780847](https://pubmed.ncbi.nlm.nih.gov/18780847/)]
19. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010 Oct;51(3-4):214-221 [FREE Full text] [doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004)] [Medline: [20558196](https://pubmed.ncbi.nlm.nih.gov/20558196/)]
20. 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](https://doi.org/10.2196/jmir.6.4.e40)] [Medline: [15631964](https://pubmed.ncbi.nlm.nih.gov/15631964/)]
21. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010;12(1):e4 [FREE Full text] [doi: [10.2196/jmir.1376](https://doi.org/10.2196/jmir.1376)] [Medline: [20164043](https://pubmed.ncbi.nlm.nih.gov/20164043/)]
22. Hardeman W, Griffin S, Johnston M, Kinmonth AL, Wareham NJ. Interventions to prevent weight gain: a systematic review of psychological models and behaviour change methods. *Int J Obes Relat Metab Disord* 2000 Feb;24(2):131-143. [Medline: [10702762](https://pubmed.ncbi.nlm.nih.gov/10702762/)]
23. Kwaliteitsinstituut voor de Gezondheidszorg CBO en Nederlands Huisartsen Genootschap. Utrecht. 2006. Multidisciplinaire richtlijn Cardiovasculair risicomangement URL: https://www.nvvc.nl/UserFiles/File/Pdf/cvrm_2006.pdf [accessed 2014-08-22] [WebCite Cache ID 6S1NPEmIW]

24. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2011 Feb 15;154(4):291. [doi: [10.7326/0003-4819-154-4-201102150-00017](https://doi.org/10.7326/0003-4819-154-4-201102150-00017)]
25. 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](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
26. Sassen B, Cornelissen V, Kiers H, Wittink H, Kok GJ, Vanhees L. Physical fitness matters more than physical activity in controlling cardiovascular disease risk factors. *Eur J Cardiovasc Prev Rehabil* 2009 Dec;16(6):677-683. [doi: [10.1097/HJR.0b013e3283312e94](https://doi.org/10.1097/HJR.0b013e3283312e94)] [Medline: [19734792](https://pubmed.ncbi.nlm.nih.gov/19734792/)]
27. Sassen B, Kok G, Schaalma H, Kiers H, Vanhees L. Cardiovascular risk profile: cross-sectional analysis of motivational determinants, physical fitness and physical activity. *BMC Public Health* 2010;10:592 [FREE Full text] [doi: [10.1186/1471-2458-10-592](https://doi.org/10.1186/1471-2458-10-592)] [Medline: [20929529](https://pubmed.ncbi.nlm.nih.gov/20929529/)]
28. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, American Heart Association, National Heart, Lung, and Blood Institute. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* 2005 Oct 25;112(17):2735-2752 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.105.169404](https://doi.org/10.1161/CIRCULATIONAHA.105.169404)] [Medline: [16157765](https://pubmed.ncbi.nlm.nih.gov/16157765/)]
29. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report. *Circulation* 2002 Dec 17;106(25):3143-3421 [FREE Full text] [Medline: [12485966](https://pubmed.ncbi.nlm.nih.gov/12485966/)]
30. Sassen B, Kok G, Vanhees L. Predictors of healthcare professionals' intention and behaviour to encourage physical activity in patients with cardiovascular risk factors. *BMC Public Health* 2011;11:246 [FREE Full text] [doi: [10.1186/1471-2458-11-246](https://doi.org/10.1186/1471-2458-11-246)] [Medline: [21504566](https://pubmed.ncbi.nlm.nih.gov/21504566/)]
31. Fishbein M, Ajzen I. Predicting and changing behavior: the reasoned action approach. New York: Psychology Press; 2010.
32. Intervention website. URL: <http://www.gezondheidsbevordering.eu/> [accessed 2014-08-22] [WebCite Cache ID 6S1P4G3fW]
33. Godin G, Bélanger-Gravel A, Eccles M, Grimshaw J. Healthcare professionals' intentions and behaviours: a systematic review of studies based on social cognitive theories. *Implement Sci* 2008;3:36 [FREE Full text] [doi: [10.1186/1748-5908-3-36](https://doi.org/10.1186/1748-5908-3-36)] [Medline: [18631386](https://pubmed.ncbi.nlm.nih.gov/18631386/)]
34. Eccles MP, Hrisos S, Francis J, Kaner EF, Dickinson HO, Beyer F, et al. Do self-reported intentions predict clinicians' behaviour: a systematic review. *Implement Sci* 2006;1:28 [FREE Full text] [doi: [10.1186/1748-5908-1-28](https://doi.org/10.1186/1748-5908-1-28)] [Medline: [17118180](https://pubmed.ncbi.nlm.nih.gov/17118180/)]
35. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychol Bull* 2006 Mar;132(2):249-268. [doi: [10.1037/0033-2909.132.2.249](https://doi.org/10.1037/0033-2909.132.2.249)] [Medline: [16536643](https://pubmed.ncbi.nlm.nih.gov/16536643/)]
36. Presseau J, Johnston M, Francis JJ, Hrisos S, Stamp E, Steen N, et al. Theory-based predictors of multiple clinician behaviors in the management of diabetes. *J Behav Med* 2014 Aug;37(4):607-620. [doi: [10.1007/s10865-013-9513-x](https://doi.org/10.1007/s10865-013-9513-x)] [Medline: [23670643](https://pubmed.ncbi.nlm.nih.gov/23670643/)]
37. de Bruin M, Viechtbauer W, Schaalma HP, Kok G, Abraham C, Hospers HJ. Standard care impact on effects of highly active antiretroviral therapy adherence interventions: A meta-analysis of randomized controlled trials. *Arch Intern Med* 2010 Feb 8;170(3):240-250. [doi: [10.1001/archinternmed.2009.536](https://doi.org/10.1001/archinternmed.2009.536)] [Medline: [20142568](https://pubmed.ncbi.nlm.nih.gov/20142568/)]
38. Légaré F, Stacey D, Brière N, Fraser K, Desroches S, Dumont S, et al. Healthcare providers' intentions to engage in an interprofessional approach to shared decision-making in home care programs: a mixed methods study. *J Interprof Care* 2013 May;27(3):214-222 [FREE Full text] [doi: [10.3109/13561820.2013.763777](https://doi.org/10.3109/13561820.2013.763777)] [Medline: [23394265](https://pubmed.ncbi.nlm.nih.gov/23394265/)]
39. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004 Feb;8(6):iii-iv, 1 [FREE Full text] [Medline: [14960256](https://pubmed.ncbi.nlm.nih.gov/14960256/)]
40. Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Acad Med* 2010 May;85(5):909-922. [doi: [10.1097/ACM.0b013e3181d6c319](https://doi.org/10.1097/ACM.0b013e3181d6c319)] [Medline: [20520049](https://pubmed.ncbi.nlm.nih.gov/20520049/)]
41. Brouwer W, Oenema A, Crutzen R, de Nooijer J, de Vries NK, Brug J. An exploration of factors related to dissemination of and exposure to internet-delivered behavior change interventions aimed at adults: a Delphi study approach. *J Med Internet Res* 2008;10(2):e10 [FREE Full text] [doi: [10.2196/jmir.956](https://doi.org/10.2196/jmir.956)] [Medline: [18417443](https://pubmed.ncbi.nlm.nih.gov/18417443/)]
42. van der Weide M, Smits J. Adoption of innovations by specialised nurses: personal, work and organisational characteristics. *Health Policy* 2004 Apr;68(1):81-92. [doi: [10.1016/j.healthpol.2003.09.007](https://doi.org/10.1016/j.healthpol.2003.09.007)] [Medline: [15033555](https://pubmed.ncbi.nlm.nih.gov/15033555/)]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
ICT: information and communication technology
PBC: perceived behavioral control

Edited by G Eysenbach; submitted 09.12.13; peer-reviewed by CA Waldron, L Lechner, S Langrial; comments to author 15.03.14; revised version received 27.04.14; accepted 22.07.14; published 21.10.14.

Please cite as:

Sassen B, Kok G, Schepers J, Vanhees L

Supporting Health Care Professionals to Improve the Processes of Shared Decision Making and Self-Management in a Web-Based Intervention: Randomized Controlled Trial

J Med Internet Res 2014;16(10):e211

URL: <http://www.jmir.org/2014/10/e211/>

doi: [10.2196/jmir.3170](https://doi.org/10.2196/jmir.3170)

PMID: [25337988](https://pubmed.ncbi.nlm.nih.gov/25337988/)

©Barbara Sassen, Gerjo Kok, Jan Schepers, Luc Vanhees. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Patients' Continuing Use of an Online Health Record: A Quantitative Evaluation of 14,000 Patient Years of Access Data

Richard G Phelps¹, FRCP (Edin), PhD; Joanne Taylor¹, MB ChB; Keith Simpson²; Jasmine Samuel¹, MB ChB; A Neil Turner¹, FRCP (Edin), PhD

¹University of Edinburgh, Renal Autoimmunity Group, Centre for Inflammation Research, Edinburgh, United Kingdom

²Glasgow Royal Infirmary, Glasgow, United Kingdom

Corresponding Author:

Richard G Phelps, FRCP (Edin), PhD

University of Edinburgh

Renal Autoimmunity Group

Centre for Inflammation Research

47 Little France Crescent

Edinburgh, EH16 4TJ

United Kingdom

Phone: 44 131 2429164

Fax: 44 131 2429164

Email: richard.phelps@ed.ac.uk

Abstract

Background: Online access to all or part of their health records is widely demanded by patients and, where provided in form of patient portals, has been substantially used by at least subgroups of patients, particularly those with chronic disease. However, little is reported regarding the longer-term patient use of patient-accessible electronic health record services, which is important in allocating resources. Renal PatientView (RPV) is an established system that gives patients with chronic kidney disease access to live test results and information about their condition and treatment. It is available in most UK renal units with up to 75% of particular patient groups registered in some centers. We have analyzed patient use out to 4 years and investigated factors associated with more persistent use.

Objective: Our aim was to investigate RPV use by patients over time from initial registration in order to understand which patients choose to access RPV and the endurance of its appeal for different patient groups.

Methods: We analyzed an anonymized extract of the database underlying RPV containing information on patient registration and events including patient access and the arrival of new blood test results or letters that patients might wish to view.

Results: At the time of the extract, there were 11,352 patients registered on RPV for 0-42 months (median 17). More than half of registrants became persistent users, logging in a median of 2.0 times each month over post-registration intervals of up to 42 months (median 18.9). Provision of assistance with first logon was strongly associated with becoming a persistent user, even at 3 years. Logons by persistent users occurred around the time of consultations/tests, strongly suggestive of patient engagement. While indices indicative of greater deprivation were the strongest determinants of non-participation, they had negligible influence on drop-out rates among established users.

Conclusions: In this mature patient portal system, a large proportion of patients made regular use of their online health records over protracted periods. The patterns and timing of use indicate strong patient interest in detailed information such as recent test results and clinic letters. Supporting patients through the first steps of establishing access to their online records is associated with much higher rates of long-term use of RPV and likely would increase use of other electronic health records provided for patients with chronic disease.

(*J Med Internet Res* 2014;16(10):e241) doi:[10.2196/jmir.3371](https://doi.org/10.2196/jmir.3371)

KEYWORDS

electronic health records; patient access to personal records; chronic renal insufficiency; utilization

Introduction

Patient-accessible electronic health records (EHRs) could have many advantages, but patient enthusiasm and use have been variable [1,2]. The more successful systems have provided information that patients or relatives need, usually in the setting of chronic disease [3-8], high criticality [9], or pertinence [10]. Variation in uptake and use has been related to social factors, and concerns have been expressed about a digital divide [11-15] and the cost in relation to utilization and health outcomes [16]. While patient use of EHRs has been evaluated in several studies, patients' continuing use over longer terms has not been reported.

Renal diseases are mostly chronic conditions and sometimes progressive, so dialysis or renal transplantation is required. Patients with progressive or end-stage renal disease are under regular specialist supervision for the rest of their lives. They may therefore have a strong incentive to become interested in the monitoring and treatment of their condition. In particular, blood tests change in important ways in kidney disease and are typically undertaken every 3-12 months in CKD patients and patients with stable transplants, and monthly in patients receiving center-based hemodialysis.

Renal PatientView (RPV) was established in 2004 to provide information to patients with chronic kidney disease (CKD) (see the Renal PatientView website for a demo [17]). RPV enables patients to see their unscreened blood test results, doctor's letters, information links, and certain other health data using a standard Web browser on any computer. There is no financial barrier or incentive to patients' use of the system. Renal units make small annual payments for access. Patients are made aware of RPV in various ways including during consultations and by local advertising. Interested patients sign a request for their health data to be sent from their treatment center to RPV and then receive the link to RPV, a user name, and password with which to access RPV from a Web browser. The first time patients log on to RPV, they are required to change their password. Every logon is recorded so it is possible to determine when and for how long patients choose to continue their use of RPV.

Patient and staff responses to RPV have been quite positive (in the case of staff, this is despite common early hesitancy) [6,18]. RPV has been offered to patients of an increasing proportion

of renal units in the United Kingdom since 2005, and by mid-2013 it was implemented in 60 of the United Kingdom's 73 renal units.

In order to better understand variation in RPV uptake and persistence of usage, we analyzed the registration and access log data from the RPV server log files collected over nearly 4 years.

Methods

Description of the Datasets

RPV data extant on September 7, 2009 (the RPV census date), were stripped of patient identifiers and transferred to a research database. Subsets of the data were extracted to R for statistical analysis (the R Foundation for Statistical Computing [19]).

The dataset comprised patient factors listed in Table 1. Deprivation scores were obtained by searching patients' postal codes against the Index of Multiple Deprivation (IMD; available for postal codes in England and Wales) and Scottish Index of Multiple Deprivation (SIMD; available for postal codes in Scotland) databases, which associate postal codes of residence with measures of deprivation determined from the UK National Census [20]. The merits of using (S)IMD measures of deprivation was discussed recently [21]; in summary, they attempt a measure of the overall deprivation experienced by people living in an area by combining 38 indicators across seven domains (income; employment; health and disability; education, skills, and training; barriers to housing and other services; crime; living environment), weighting in particular income and employment. Deciles of rank deprivation were analyzed in preference to raw scores to reduce the number of levels for comparisons and mitigate any effects of the slightly different weighting used in deriving IMD and SIMD scores. Focus on the extremes of the deprivation range was achieved by further consolidation of the rank scores into three levels as described in Table 1.

Center factors were defined to capture possible influences from the centers providing renal services to patients. Centers were classified according to the date RPV was first offered to patients, the proportion of the Renal Replacement Therapy (RRT) population recruited to use RPV, and their provision or not of extra support to patients registering with RPV for the first time (assisted start).

Table 1. Patient factors.

Patient factors		Notes
Age in years	<18, 18-34, 35-54, 55-74, >75	
Gender	M/F	
Treatment		
	Hospital HD ^a	Includes most dependent patients
	Home HD/PD ^b	
	Transplant	
	Not on RRT ^c	
Deprivation^d		
By decile	1-10	1 is most deprived
By three groups		
	High	Deciles 1 and 2
	Middle	Deciles 3-8
	Low	Deciles 9 and 10
Access log data	Dates and times of every logon	
Blood test results	Sample dates and values	
Center	Unit code	UK renal registry code of treatment center

^aHD: hemodialysis.

^bPD: peritoneal dialysis.

^cRRT: renal replacement therapy, so HD, PD, or transplant.

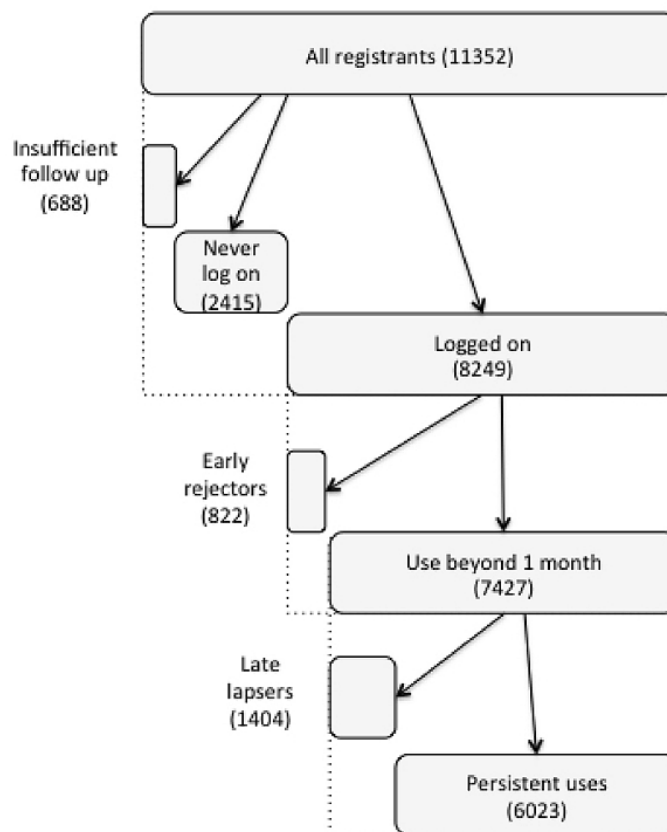
^dDeprivation measures: see text.

Classification of Users by Logon Activity

Figure 1 shows the approach taken to classify registrants according to their completion of the initial logon and password change procedures, their making logons during the first month post-registration through making logons right up to the time the RPV census was taken (persistent users). Registrants were classified into the following groups: (1) insufficient interval of follow up- patients enrolled within 6 months of the RPV census date were excluded from classification on the basis of logon activity as insufficient time had passed to determine any pattern, (2) patients that never log on, (3) early lapsed, that is, patients who log on only during the first month post registration, (4) late

lapsed, that is, patients who at the time of census had not logged on for 6 months and for whom there had been at least 2 sets of blood result uploads (patients were therefore not deceased and had some incentive to log on), and (5) persistent users, that is, all other patients.

For some studies, patients were also classified by treatment in order to enable comparison with an appropriate overall UK population. This enabled comparison of adult RPV registrants in receipt of renal replacement therapy with the overall UK adult population in receipt of renal replacement therapy at the time of the RPV census as recorded at the UK Renal Registry report 2009 [22].

Figure 1. Classification of RPV users by logon analysis.

Exploration of Factors Possibly Associated With Registration and Usage

Statistical analysis used the R framework [19]. Logistic regression used the glm program of the R stats package. The approach was to initially include all the factors listed above, then iteratively remove less significant factors and test the reduction in variance explained by the simplified model (analysis of variance) to decide their removal.

Persistence of Use

Survival (of RPV use) analysis employed the survplot program [23]. Use was viewed as beginning on date of first logon and ending on date of last logon. Use durations were right-censored where last logon dates were judged consistent with continuing use of the system: specifically, last recorded logon was either within 6 months of the RPV census date, or earlier but not succeeded by at least 2 results events (suggesting infrequent follow-up). The last provision also right-censored patients who had been users of RPV up to death.

Analysis of Logon Activity

The overall intensity of use was computed by counting, for each patient who completed at least 1 logon, the total number of logons made and dividing by the number of months between registration and census to compute logons per month. The distribution of the logarithm of logon activity across the patient population was approximately normal (see Results), enabling patients to be classified by the quartile (Q1-Q4) that their individual log (logons per month) fell in the overall log (logons per month) distribution.

The timing of logon events in relation to test results was explored first with a plot devised to show the activity of individual patients as series of colored dots along a horizontal timeline. The activity of approximately 100 patients could be plotted without obscuring individual patient events, so plots show random selections of patients from appropriate subgroups ordered by the duration of logon activity. The overall tendency of logon events to occur close in time to test result events was explored by computing for every logon event the interval to the closest result event and constructing a histogram to show the proportion of logons by interval to closest test date. Significance was assessed by comparing histograms made with randomly shuffled logon times or test times.

Results

Which Patients Register to See Their Records?

At the time of the RPV census, 11,352 patients had been registered from 37 of the United Kingdom's 73 adult centers. The characteristics of the 11,352 registrants are shown in Table 2. Their age and sex distribution broadly parallels the age range and male propensity of renal disease [22]. To assess factors that might be associated with registration for RPV, the characteristics of registrants in receipt of renal replacement therapies (n=6646) was compared with that of all UK patients in receipt of renal replacement therapies (N=44,649). Overall, 14.8% of UK adult RRT patients were registered with RPV, but there were significant differences in the proportion registered by age, treatment, and deprivation (Table 3). The proportion registered in the 18-34 and 35-54 age groups were almost twice that of

the over 75 age group, and the proportion among patients with transplants or on home-based dialysis (peritoneal dialysis or home hemodialysis) was about 70% greater than of that of hospital hemodialysis patients. The proportion registered increased almost linearly with increasing rank deprivation score (ie, with lower deprivation) (Figure 2). Patients residing at postal

codes associated with the lowest levels of deprivation (rank 10) were 2.4 (England and Wales) to 3.2 (Scotland) times more likely to be registered for RPV than those at postal codes associated with the highest deprivation (rank 1). There were no significant differences by gender.

Table 2. Description of the entire RPV-registered population (N=11,352).

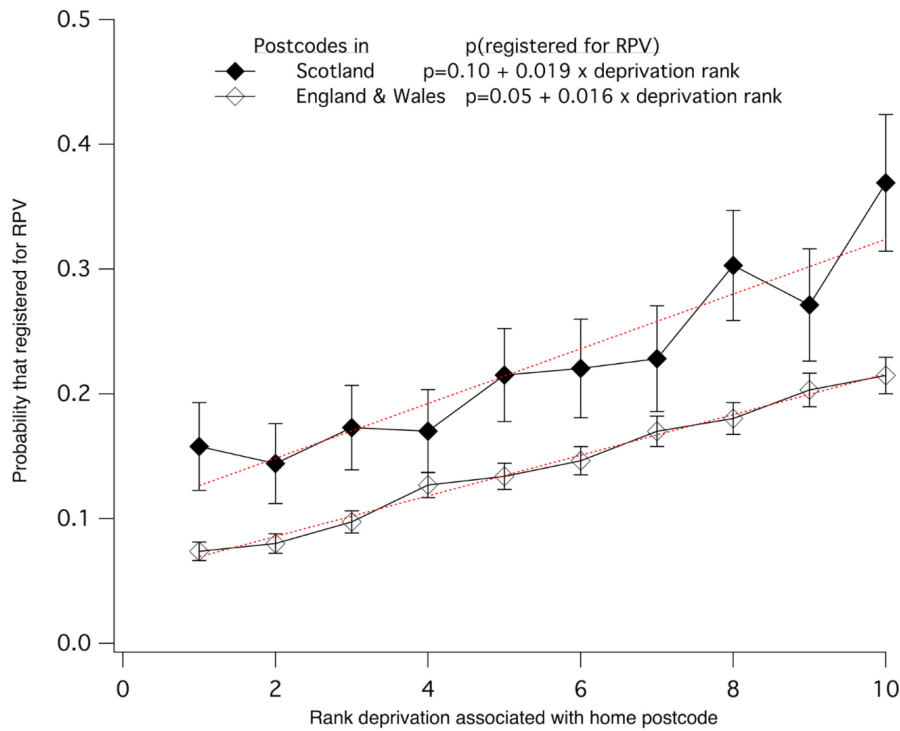
Characteristic	Proportion, %
Age in years	
<18	1.6
18-34	11.6
35-54	37.9
55-74	38.9
>75	10.0
Gender	
Male	59.9
Treatment	
Hospital HD	19.4
Home HD/PD	6.7
Transplant	32.8
Not on RRT	41.0

Table 3. Description of the RRT-dependent subgroup (n=6646) of the RPV-registered population compared with the overall UK RRT population by age, gender, and treatment (N=44,649)^a.

Characteristic	Proportion registered for RPV of all UK RRT patients, %	P value
Age in years		
18-34	16.1	<i>P</i> <.001
35-54	16.6	
55-74	14.8	
>75	8.4	
Gender		
Male	15	Not significant
Female	14.6	
Treatment		
Hospital HD	10.5	<i>P</i> <.001
PD	17.7	
Transplant	16.6	
Overall		
	14.8	

^aSignificance of comparisons estimated by Pearson's chi-square test statistic.

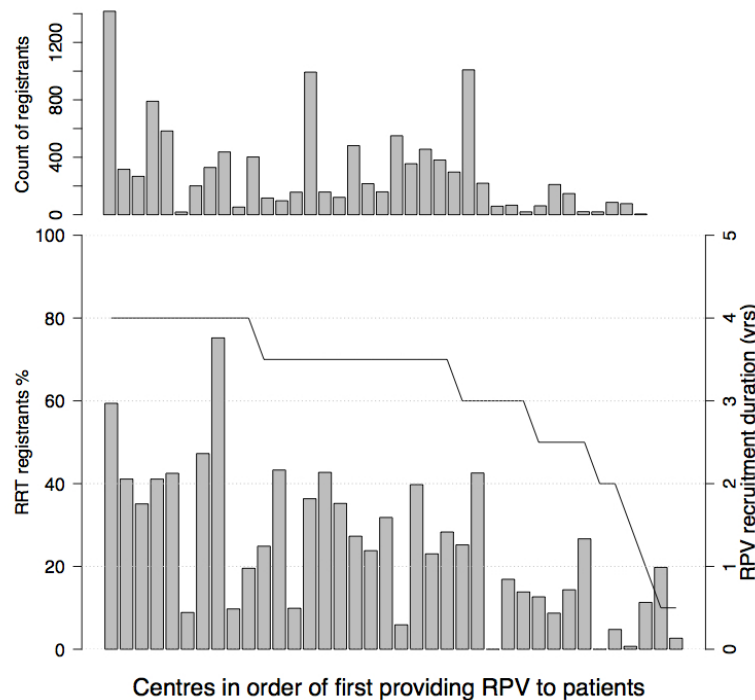
Figure 2. RPV registration by patient deprivation. The proportion of adult RRT patients registered for RPV is shown by patient rank deprivation (1-10, 10 is least deprived) for patients with postal codes in Scotland (filled diamonds) or England & Wales (open diamonds).



The penetration (proportion of adult RRT patients registered) by center ranged from 6-75% (median 32) for centers enrolling at least 6 months. The variation was only in part explained by differences in the duration of recruitment effort (6-48 months) (Figure 3), so very likely it was influenced by centers' recruitment practices. One very significant difference in

recruitment practice was identified (see below), but formal definition of the influential factors is a goal of ongoing investigation. It is noteworthy, however, that many of the centers with higher recruitment have enthusiastic local proponents of RPV within their clinical teams.

Figure 3. RPV registration by center. Number of registrants (top chart) and proportion (%) of available RRT patients registered for RPV is shown by renal center, ordered by the start date of patient enrollment. The duration of active RPV enrollment is superimposed on the bottom chart; duration ranged from 0.5-4 years.



Completing First Logon

About one-fifth (23.20%, 2634/11,352) of registered patients never logged in despite having signed up and having been sent details. Logistic regression indicated that the strongest registrant-specific effects (up to 2-fold) were age and treatment group (Table 4). Registrants in middle age were more likely to complete first logon than younger (<34) and older (>75) registrants, and registrants with a transplant were more likely to log on than patients on hemodialysis. The complex influence of age was further studied avoiding arbitrary age grouping by using a general additive model and non-parametric smoother (Figure 4B). This revealed that the small number of very young registrants were the most likely of all to complete first logon, presumably represented by their parents.

The effect of deprivation was also very significant. Registrants from addresses associated with greatest deprivation were more likely to not complete first logon (OR 1.24, 95% CI 1.08-1.42), whereas registrants with addresses associated with low levels of deprivation were less likely to not complete first logon (OR 0.79, 95% CI 0.65-0.97), both compared with registrants of middle rank deprivation, indicating a negative influence of deprivation beyond that on becoming registered, as described in other chronic disease populations [11,12].

Patients at centers that began offering RPV in the 2 years prior to the census were twice as likely to not complete first logon (OR 2.04, 95% CI 1.41-2.93) as patients at centers that had been offering RPV for more than 2 years, which seemed greater than the lag to be expected for recently enrolled centers to build up recruitment. This is a subject of continuing investigation.

Table 4. Odds of not persisting with RPV use at proposed hurdles^{a,b}.

	Odds not completing initial logon (N=9552), OR (CI) ^c	Odds lapse early (N=822 vs 7427), OR (CI)	Odds lapse late (N=1401 vs 6023), OR (CI)
Age (compared with age 35-54)			
<18	1.81 (1.27-2.55)	1.02 (NS)	0.76 (NS)
18-34	1.32 (1.12-1.56)	1.21 (NS)	1.03 (NS)
55-75	0.98 (NS)	1.02 (NS)	1.11 (NS)
>75	1.41 (1.19-1.66)	1.46 (1.12-1.90)	1.58 (1.27-1.96)
Treatment (compared with hospital HD)			
Pre RRT	1.0 (NS)	0.83 (NS)	0.37 (0.31-0.44)
Home HD/PD	0.88 (NS)	0.96 (NS)	0.68 (0.53-0.86)
Transplant	0.60 (0.52-0.69)	0.61 (0.49-0.75)	0.45 (0.38-0.52)
Deprivation (compared with middle)			
Greatest	1.24 (1.08-1.42)	1.73 (1.40-2.13)	—
Least	0.79 (0.70-0.89)	0.79 (0.65-0.96)	—
Unit offering RPV <2 years	2.04 (1.41-2.93)	—	—
Unit registration rate ^d in Q3 (Q2 also significant)	0.79 (0.65-0.97)	—	—
Unit offering routine assisted start	0.31 (0.21-0.46)	—	—

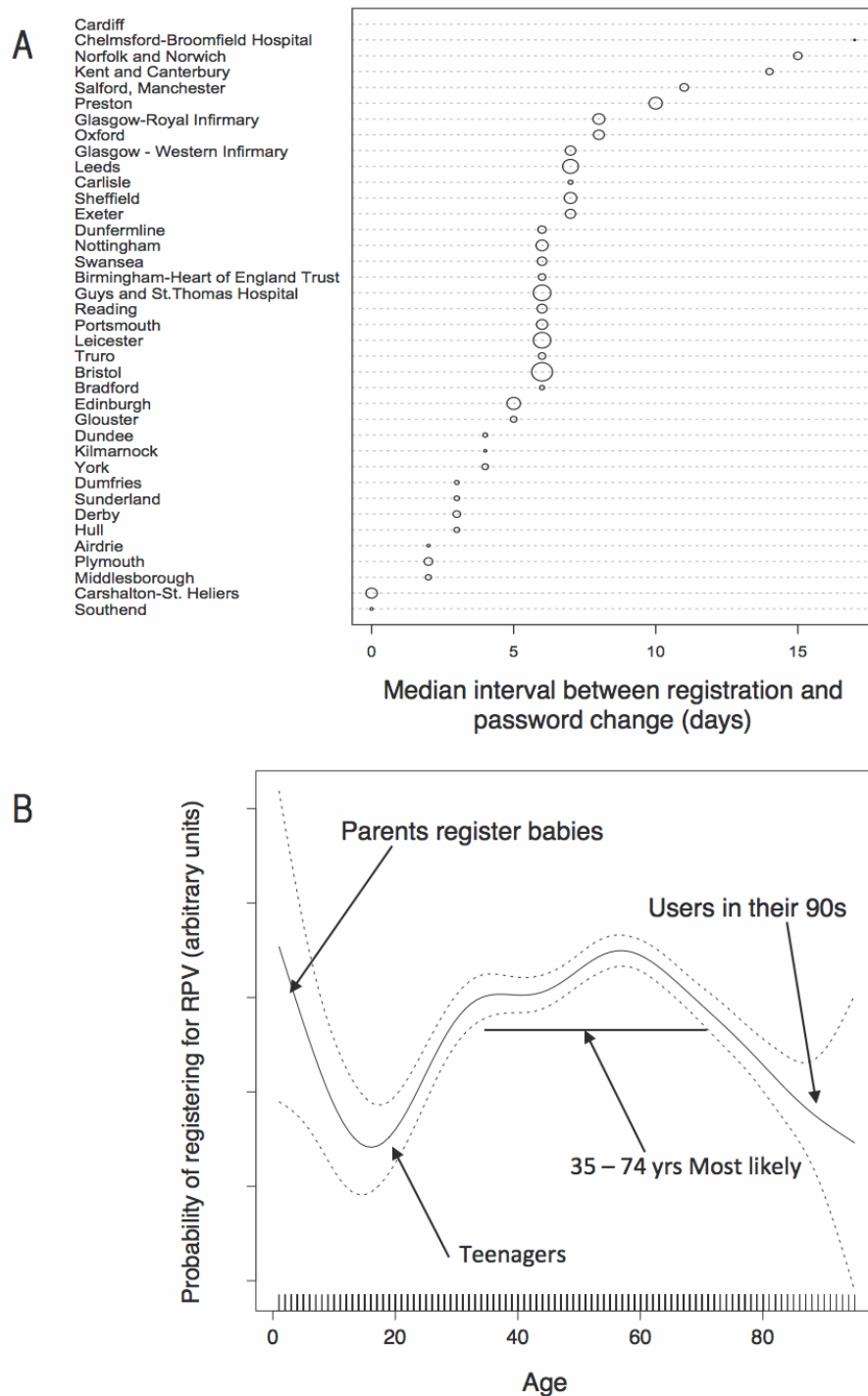
^aSummaries of models obtained by logistic regression for the likelihood that patients choose not to persist with use of RPV at the three decision points: (1) choosing not to complete first logon, (2) having made an initial logon choosing not to logon on again beyond 1 month (lapse early), and (3) discontinuing logons at some later time (lapse late).

^bOdds ratios are shown for the influential factors followed in brackets with 95% confidence intervals or NS if the interval spans 1.0. Factors marked with a dash (—) were removed because of insignificant effects in the indicated model. Gender had insignificant effects in all models.

^cAnalysis restricted to the 9552 (of 11,352) registrants with complete data. The major reason for exclusion was missing treatment type as a result of this parameter not being recorded by one major center (1009 registrants). Alternative analysis including these 1009 registrants and excluding treatment as a factor did not for any other factor change the assignment of significance and only slightly altered the ORs.

^dCenters were grouped in quartiles of percentage enrollment of RRT patients as a measure of a center's effectiveness in recruiting patients to RPV.

Figure 4. Interval from registration and likelihood of completing first logon. A: the median interval in days between registration and initial patient logon by renal center. The size of markers is proportional to total number of registrants at each center. Two centers are remarkable for completing the process within a day for most patients (87% of patients at the larger center). B: Probability of completing first logon by age. A logistic (completed first logon = T/F) general additive model (non-parametric) was constructed using the mgcv package to model the likelihood of patients completing first logon by age at time of registration without assumption as to the shape of any relationship.



Interval Between Registration and First Patient Logon

The interval between registration and first logon was investigated because it was thought likely to be a marker of patient enthusiasm. However, the interval was dominantly influenced by center (Figure 4A). Most striking was the very short interval (<1 day) at two centers (Carshalton and Southend)

leading to the discovery that these centers had elected to have an administrator help patients complete their first logon immediately after registering, in contrast to the usual practice of sending logon instructions by mail. Unsurprisingly, patients of these 2 centers were unlikely to not complete first logon (OR 0.31, 95% CI 0.21-0.46) when compared to patients of all other

centers. More importantly, this practice also influenced subsequent logon behaviors (see below).

Persistence of Use After First Logon

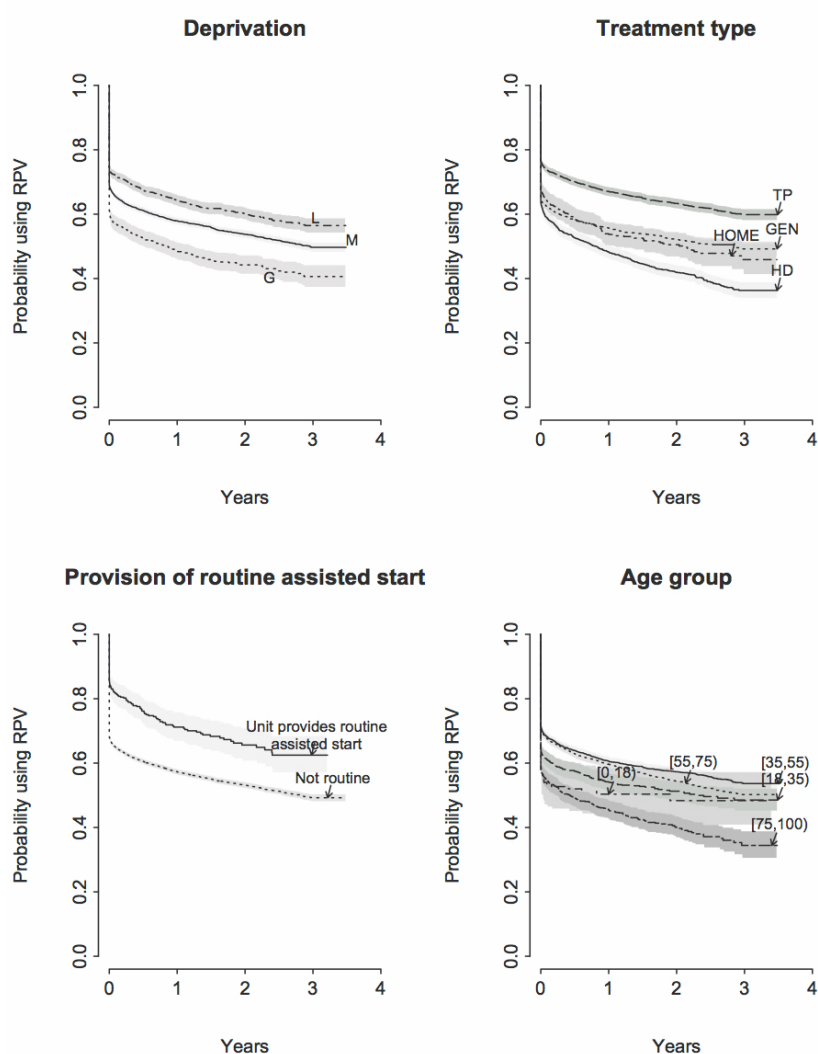
There was substantial variation in both the frequency and persistence of RPV logon activity subsequent to first logon. As some registrants had been registered only shortly before the RPV census such that their logon activity was too brief to be assessed, analysis of subsequent logon activity was restricted to the 8249 registrants who completed first logon more than 3 months prior to the census. Three broad patterns were distinguished (Figure 1); 822/8249 (9.96%) made no logon beyond the first month suggesting they were disinclined to continue engagement (early lapse). Of the 8249 (90%) who made use of RPV beyond 1 month, most (6023/8249, 73.01%) made continued use up to the census date (persistent users), but 1404 (16.98%) were judged to have lapsed (perhaps wrongly, see Discussion) in their use of RPV because at least 6 months had passed and at least 2 sets of test results had been uploaded since last logon (continued arrival of test results confirming that they were still alive). The 8249 registrants who made at least one logon went on to make a total of 1-914 (median 14, mean 42.7) logons over 0-42 months. Among those classified as lapsing and persistent users, the median number of logons was 4 and 26 respectively.

Factors that might influence subsequent logon behavior were investigated first by logistic regression. Early lapse was associated with age over 75 and greater deprivation and was less likely in transplant recipients (Table 4). Late lapsing was associated with age over 75 and treatment by hospital hemodialysis. Interestingly, deprivation was not a significant factor in this cohort of established RPV users.

Survival analysis was used to better understand lapses in use of RPV over time, plotting probability of continuing use by time for users partitioned by relevant factors (Figure 5). Overall, the probability of continuing use was 0.61 (95% CI 0.602-0.621) at 6 months falling gradually by about 0.04 per year (6 months 0.61, 12 months 0.58, 18 months 0.56, 24 months 0.54, 30 months 0.52). Interestingly, while greater deprivation was associated with substantially reduced probability of continuing RPV use at 6 months, as might be expected [14], the subsequent rates of attrition were similar in all three deprivation groups. In contrast, the greater early attrition in the elderly and hemodialysis treated was followed by a continuing higher rate of attrition than in the young and those with transplants.

The substantially greater early use observed in users at centers offering assisted starts was succeeded by a slightly greater rate of attrition out to 2.5 years leading to some convergence of the survival curves, but the beneficial influence remained highly significant.

Figure 5. Probability of persisting use of RPV at intervals after registration shown as survival (of RPV use) classified as shown. Shading indicates 95% confidence limits.



Patterns of Use Over Time and in Relation to Test Results

An immediate peak of very frequent logons was observed, which settled to a slowly declining average rate of logons. This concealed a highly variable rate of logons for individual patients, with periods of intense activity associated with periods of more frequent test results. These periods are likely to be times of medical uncertainty; for instance, some could be around transplantation or other hospital admissions.

Frequency of use overall was assessed from the number of logons made by patients divided by the interval between first and last logon. This analysis was limited to the 5808 users who had made at least three logons over at least 3 months after first logon; 87.55% (5085/5808) were persistent users and 12.45% (723/5808) late lapsed. The median number of logons per month was 1.8 (2 for persistent users, 0.94 for lapsed) with a markedly skewed distribution (range 0.08-84) (Figure 6A). Subsequent comparisons were made of the natural logarithm of logons per month, which was distributed near normally (Figure 6B), categorized by quartile (Q1-Q4 where Q4 exhibits the heaviest use). The proportion of male and female users with Q1-Q4 logon frequency was very even, but with respect to age, users in the

two younger age groups had a predominance of higher frequency logon activity (Q3 and Q4) and users in the oldest age group had a predominance of lower level activity (Q1 and Q2) (Figure 6C). In contrast with the negative influence of greater deprivation on registration and initial logon, within this cohort of established users, patients with addresses associated with greater deprivation exhibited a slight predominance of higher logon activity (Figure 6D). With respect to treatment type (Figure 6E), users on home-based dialysis included the highest proportion of higher-level users, which suggests patient self-monitoring. Higher-level use was also prevalent among users on hospital hemodialysis, a population that in most centers has frequent blood test monitoring (commonly monthly), providing frequent incentive to check results.

The patients exhibiting the most frequent logon activity (top 5%) were also analyzed separately. None of the factors available was associated with this extreme pattern of use.

Most logon events occurred on weekdays and between 08:00-22:00; in particular, Sunday logon was not popular (Figure 6G and H). No differences in the times of logon events was observed by age, sex, treatment group, or deprivation.

The time of logons to RPV was also examined in relation to the most proximate blood tests because surveys had suggested that accessing blood test results was a major reason for using RPV. First logon events were plotted against interval since first logon to visualize the patterns of logon activity. Examples are shown in Figure 7. There was clear variation in patient logon activity with the most intense activity typically occurring over short periods during which it is likely clinical circumstances increased the need for frequent updates. Most events occurred in the week after new blood test results became available when logons were likely performed to get the results. This was investigated further

by determining for all logon events the interval from the most proximate new test result date: 75% of logons occurred within 2 weeks of new results becoming available (Figure 7C). The most likely interval was 0-1 day (21.6% of all logons), followed by 1-2 and 2-3 days (0-3 together 45% of logons), but surprisingly the 4th most common interval was the 24 hours before blood tests were taken, when results could not possibly be available. Indeed, 10.8% of all logons occurred in the 3 days leading up to a blood test, suggesting that registrants are reviewing their results before a clinical encounter.

Figure 6. Logon activity. A-F: Patients classified by quartile of log (count of logons per month) (defined in panels A and B as described in the Methods), and comparisons made of the partitioning of Q1-Q4 activity patients by the factors shown. The even spread by gender contrasts with that by age (C) and treatment type (E). G-H: Histograms of counts of logons made by day of week (G) and time of day (H).

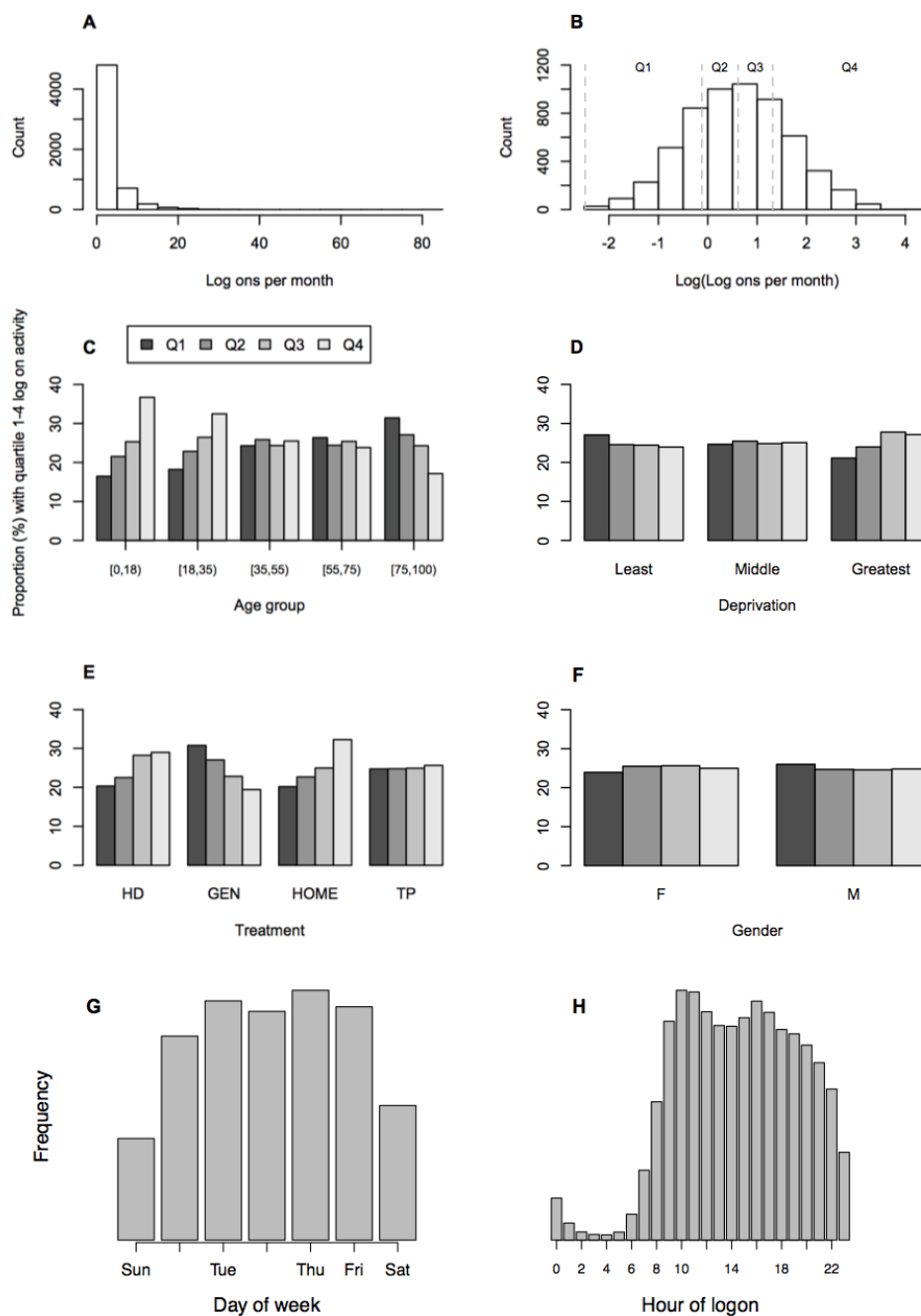
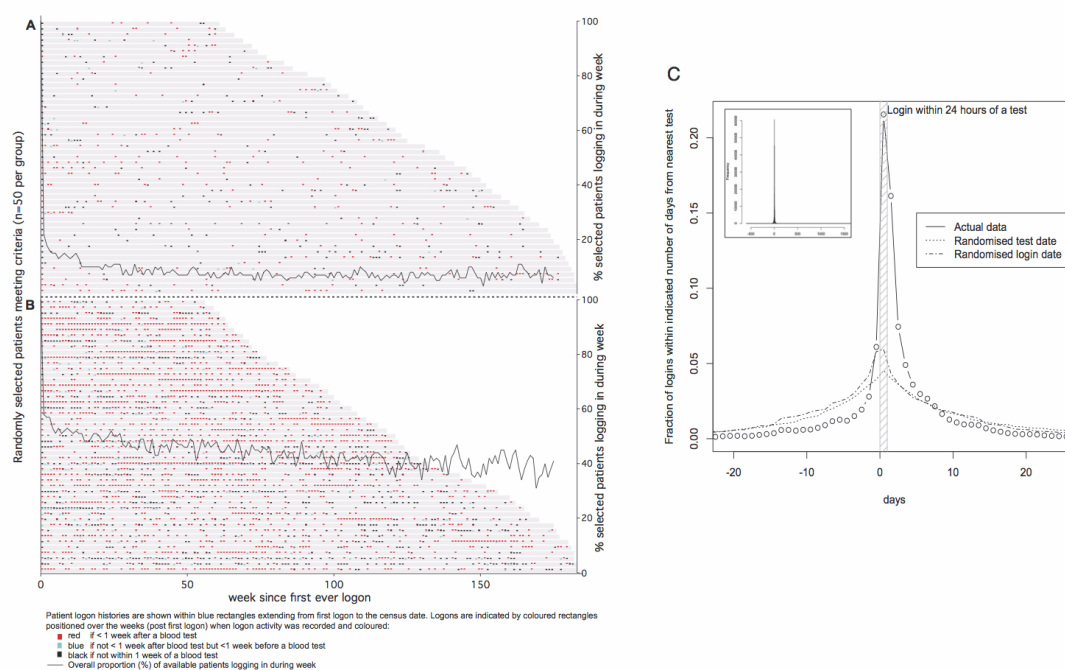


Figure 7. Relationship between times of patient logons and blood test results. A, B: Individual logon events shown by interval (in weeks) since first logon for randomly selected lower (A) and higher (B) frequency transplant recipient users where each selected patient's activity is depicted by a horizontal series of colored dots. Most logon events are red indicating they occurred less than 1 week after a new blood result event. Superimposed is the proportion (% on the right hand axis) of patients in the respective activity groups that log in each week post first logon. C: Histogram (1 day bins) showing the predominance of logons occurring 1-3 days after tests is strikingly different from that observed with randomly shuffled data.



Discussion

High Rate of Persistent Use

It is striking and encouraging for advocates of patient-accessible EHRs that over half of registrants became persistent users (median of 2 logons per month) with only a small proportion (about 4% per year) discontinuing regular use out to 3.5 years. More persistent use was expected as CKD is a chronic condition, but the rate is much higher than the 26-33% reported in other studies that have assessed use by patients with chronic diseases (eg, [7,8]). RPV was developed by clinicians with close involvement of patients and iteratively improved over years to meet patients' needs, but other systems make similar claims. Clearly RPV is providing services with an enduring appeal to users, and the results suggest that key among these is the timely provision of blood test results.

Access to Timely Blood Test Results is a Key Driver to Renal PatientView Use

Blood test results are important in the continuing management of CKD informing, for example, the need for and effectiveness of dietary prescriptions as well as the progression of kidney failure or transplant performance, and many patients take a close interest encouraged by their doctors. Provision of timely blood test results was a core design goal for RPV and was identified as a key attraction for users in a small survey [6]. The current data provide strong support for the importance of this service demonstrating that 45% of all logons occur in the 3 days after a blood test was taken, presumably to look up the results. As well as indicating that access to results is what patients want, the finding is also indicative of a patient group that is engaged

with their treatment—why otherwise look up the results and try to interpret their meaning? Further support comes from the observation that almost 11% of logons took place shortly before clinical encounters, suggestive of engaged patients wishing to be current with their results before potential opportunities to discuss their management with clinicians. Certainly clinicians report that consultations with some patients have changed as a result of patients' access to RPV. An important question for ongoing study is the influence, if any, on the achievement of treatment goals as a result of greater patient involvement.

Frequency and Timing of Use

Analyzed as a group, after an early phase of frequent logons, during which patients may be gaining familiarity with the system and learning from the information links provided, the frequency of logons by persistent users settles to a median of 2 logons per month with most logons during usual waking hours and the working week (Figure 6F, G). This is approximately twice the rate at which new results become available for most patients. Analysis at the individual patient level reveals considerable variation including some patients making very frequent logons (several per day), but in most cases this was observed to be short lived (a few weeks or months), possibly reflecting periods of greater concern, such as around the time of transplantation. Greater understanding will require a formal qualitative analysis, which is underway.

Initial Patient Support Increases Persistent Patient Use

It is well recognized that initial authentication of new users is a substantial barrier for some patients. In the current data, 20% of patients never made a single logon despite having completed

a registration form, returned it to the center administration, and become registered. The first logon procedure for new RPV users is explained in the registration letter and requires navigation of an internet browser to the RPV website, entering the temporary credentials and creating a new password. An important observation from the current data is that providing additional support targeted just at this first logon hurdle has a profound and long-lasting positive influence on patient use. Two centers adopted the practice of providing initial support in the form of face-to-face or over-the-phone guidance through the first logon procedure, with about a 20% increase in the probability of patients' continuing use of RPV at 3 years (Figure 5, bottom left). Most centers give little to no active instruction at present. It may be that sufficient support could be provided more broadly without unmanageable cost implications. This will become clearer if the practice becomes more widespread as is being encouraged by the current guidance to units providing RPV [24]. It must be noted that this was a serendipitous retrospective finding, so further study is required. However, the results support a compelling case to better understand the practical issues faced by patients in accessing online records and the effectiveness of support strategies, both of which we are addressing in current research.

Mixed Inclusiveness of Access

Inclusiveness is a concern in digital service provision [11,12,15,25,26]. It is remarkable that we observed no sex difference in any of the measures studied, as most studies show that women are more likely to seek medical advice and information both in person and online. Age was influential—RPV users are slightly younger than non-users (3 years on average)—but this small difference may be reducing

as it was 7 years in an earlier survey of RPV users, admittedly comparing results obtained by different methods (survey versus log data). It is possible that logon sharing decreased sex and age differences. RPV encourages patients to share logons and many do. Up to 40% have shared logons in small surveys, and some are dependent on others to log on for them.

Level of deprivation was strongly associated with RPV registration and initial use. This is a new observation with regard to RPV but was anticipated from the results of many studies of access to EHRs and health care information on the Internet [11,12,25,26]. However the frequency of use of RPV is actually higher and the rate of drop-out after first logon is lower among users of greater deprivation. The explanation for this difference could be trivial (deprivation by postal code is imperfect) or very significant for best design of approaches to recruitment and initial support that are effective across the patient population. Digital exclusion is identified by the Department of Health in the United Kingdom as an area of high current concern [27]. Follow-up studies will address this and also determine whether the deprivation gradient is reducing with time (as it seems to be for age). Other means of access to RPV (digital TV, smartphones) are also under consideration.

Conclusions

Renal PatientView attracts strong and sustained use by many renal patients in the United Kingdom. The current data indicate that prevalent center and patient factors underlie considerable variation in RPV uptake and use. These require further study, but it is already clear that centers should consider reviewing their approach to recruiting and supporting patients and possibly identify patients that are not using RPV post-registration to offer targeted help.

Acknowledgments

This investigation was partly supported by the SUSTAINS (Support Users to Access Information and Services) project to enhance patient access to electronic health records in Europe.

Conflicts of Interest

None declared

References

1. Archer N, Fevrier-Thomas U, Lokker C, McKibbin KA, Straus SE. Personal health records: a scoping review. *J Am Med Assoc* 2011;18(4):515-522 [FREE Full text] [doi: [10.1136/amiajnl-2011-000105](https://doi.org/10.1136/amiajnl-2011-000105)] [Medline: [21672914](https://pubmed.ncbi.nlm.nih.gov/21672914/)]
2. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. *BMJ* 2010;341:c5814 [FREE Full text] [Medline: [21081595](https://pubmed.ncbi.nlm.nih.gov/21081595/)]
3. Shaw RJ, Ferranti J. Patient-provider internet portals--patient outcomes and use. *Comput Inform Nurs* 2011 Dec;29(12):714-8; quiz 719. [doi: [10.1097/NCN.0b013e318224b597](https://doi.org/10.1097/NCN.0b013e318224b597)] [Medline: [21697705](https://pubmed.ncbi.nlm.nih.gov/21697705/)]
4. Winkelman WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. *J Am Med Inform Assoc* 2005;12(3):306-314 [FREE Full text] [doi: [10.1197/jamia.M1712](https://doi.org/10.1197/jamia.M1712)] [Medline: [15684128](https://pubmed.ncbi.nlm.nih.gov/15684128/)]
5. Osborn CY, Mayberry LS, Mulvaney SA, Hess R. Patient web portals to improve diabetes outcomes: a systematic review. *Curr Diab Rep* 2010 Dec;10(6):422-435 [FREE Full text] [doi: [10.1007/s11892-010-0151-1](https://doi.org/10.1007/s11892-010-0151-1)] [Medline: [20890688](https://pubmed.ncbi.nlm.nih.gov/20890688/)]
6. Bartlett C, Simpson K, Turner AN. Patient access to complex chronic disease records on the Internet. *BMC Med Inform Decis Mak* 2012;12:87 [FREE Full text] [doi: [10.1186/1472-6947-12-87](https://doi.org/10.1186/1472-6947-12-87)] [Medline: [22867441](https://pubmed.ncbi.nlm.nih.gov/22867441/)]

7. Weppner WG, Ralston JD, Koepsell TD, Grothaus LC, Reid RJ, Jordan L, et al. Use of a shared medical record with secure messaging by older patients with diabetes. *Diabetes Care* 2010 Nov;33(11):2314-2319 [FREE Full text] [doi: [10.2337/dc10-1124](https://doi.org/10.2337/dc10-1124)] [Medline: [20739686](https://pubmed.ncbi.nlm.nih.gov/20739686/)]
8. Wagner PJ, Dias J, Howard S, Kintziger KW, Hudson MF, Seol YH, et al. Personal health records and hypertension control: a randomized trial. *J Am Med Inform Assoc* 2012;19(4):626-634 [FREE Full text] [doi: [10.1136/amiajnl-2011-000349](https://doi.org/10.1136/amiajnl-2011-000349)] [Medline: [22234404](https://pubmed.ncbi.nlm.nih.gov/22234404/)]
9. Burke RP, Rossi AF, Wilner BR, Hannan RL, Zabinsky JA, White JA. Transforming patient and family access to medical information: utilisation patterns of a patient-accessible electronic health record. *Cardiol Young* 2010 Oct;20(5):477-484. [doi: [10.1017/S1047951110000363](https://doi.org/10.1017/S1047951110000363)] [Medline: [20456816](https://pubmed.ncbi.nlm.nih.gov/20456816/)]
10. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. *Ann Intern Med* 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: [10.7326/0003-4819-155-12-201112200-00003](https://doi.org/10.7326/0003-4819-155-12-201112200-00003)] [Medline: [22184688](https://pubmed.ncbi.nlm.nih.gov/22184688/)]
11. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. *Arch Intern Med* 2011 Mar 28;171(6):568-574. [doi: [10.1001/archinternmed.2011.34](https://doi.org/10.1001/archinternmed.2011.34)] [Medline: [21444847](https://pubmed.ncbi.nlm.nih.gov/21444847/)]
12. 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](https://doi.org/10.1136/jamia.2010.006015)] [Medline: [21262921](https://pubmed.ncbi.nlm.nih.gov/21262921/)]
13. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. *J Gen Intern Med* 2011 Oct;26(10):1117-1123 [FREE Full text] [doi: [10.1007/s11606-011-1749-y](https://doi.org/10.1007/s11606-011-1749-y)] [Medline: [21647748](https://pubmed.ncbi.nlm.nih.gov/21647748/)]
14. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. *J Am Med Inform Assoc* 2011 Dec;18 Suppl 1:i8-12 [FREE Full text] [doi: [10.1136/amiajnl-2011-000473](https://doi.org/10.1136/amiajnl-2011-000473)] [Medline: [22071530](https://pubmed.ncbi.nlm.nih.gov/22071530/)]
15. 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](https://doi.org/10.1197/jamia.M1672)] [Medline: [15561786](https://pubmed.ncbi.nlm.nih.gov/15561786/)]
16. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Ann Intern Med* 2013 Nov 19;159(10):677-687. [doi: [10.7326/0003-4819-159-10-201311190-00006](https://doi.org/10.7326/0003-4819-159-10-201311190-00006)] [Medline: [24247673](https://pubmed.ncbi.nlm.nih.gov/24247673/)]
17. Renal PatientView demo site. URL: <https://www.patientview.org/> [accessed 2014-02-10] [WebCite Cache ID 6NHjxqIjf]
18. Mukoro F. A system which provides patients online access to their test results. Final Evaluation Report.: NHS Kidney Care; 2012. URL: <http://www.nhs.uk/8582.aspx> [accessed 2014-02-10] [WebCite Cache ID 6NHmXaBb6]
19. The R Project for Statistical Computing. URL: <http://www.r-project.org/> [accessed 2014-02-10] [WebCite Cache ID 6NHmENonq]
20. UK Data Service. Deprivation data. URL: <http://census.ukdataservice.ac.uk/get-data/related/deprivation.aspx> [accessed 2014-10-07] [WebCite Cache ID 6T8riwpoa]
21. Payne RA, Abel GA. UK indices of multiple deprivation - a way to make comparisons across constituent countries easier. *Health Stat Q* 2012(53):22-37. [Medline: [22410690](https://pubmed.ncbi.nlm.nih.gov/22410690/)]
22. UK Renal Registry Report 2009. URL: <http://www.renalreg.com/Reports/2009.html> [accessed 2014-09-17] [WebCite Cache ID 6SetFT0et]
23. Eklund A. Survplot. URL: <http://www.cbs.dtu.dk/~eklund/survplot/> [accessed 2014-02-10] [WebCite Cache ID 6NHmqYDIQ]
24. RIXG. Recruitment Toolkit. URL: <http://www.rixg.org/developer/pages/how-to/recruitment-toolkit.php> [accessed 2014-02-10] [WebCite Cache ID 6NHnQaacf]
25. 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](https://doi.org/10.1177/1460458211414843)] [Medline: [21937464](https://pubmed.ncbi.nlm.nih.gov/21937464/)]
26. 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](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]
27. Government UK, Health DO. The Power of Information. 2012 May. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213689/dh_134205.pdf [accessed 2014-02-10] [WebCite Cache ID 6NHn0OrKz]
28. ccc. ccc. ccc URL: <http://census.ukdataservice.ac.uk/get-data/related/deprivation.aspx> [accessed 2014-10-07] [WebCite Cache ID 6T8rr5ymA]

Abbreviations

- CKD:** chronic kidney disease (irreversible, often progressive kidney damage)
- HD:** hemodialysis
- EHR:** electronic health record

PD: peritoneal dialysis

RPV: Renal PatientView

RRT: renal replacement therapy (such as dialysis or kidney transplant)

Edited by G Eysenbach; submitted 03.03.14; peer-reviewed by D O'Donoghue, K Nazi; comments to author 17.07.14; revised version received 28.08.14; accepted 22.09.14; published 24.10.14.

Please cite as:

Phelps RG, Taylor J, Simpson K, Samuel J, Turner AN

Patients' Continuing Use of an Online Health Record: A Quantitative Evaluation of 14,000 Patient Years of Access Data

J Med Internet Res 2014;16(10):e241

URL: <http://www.jmir.org/2014/10/e241/>

doi: [10.2196/jmir.3371](https://doi.org/10.2196/jmir.3371)

PMID: [25344621](https://pubmed.ncbi.nlm.nih.gov/25344621/)

©Richard G Phelps, Joanne Taylor, Keith Simpson, Jasmine Samuel, A Neil Turner. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 24.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Promoting Health in Virtual Worlds: Lessons From Second Life

Reima Suomi¹, DSc; Matti Mäntymäki¹, DSc; Sari Söderlund², MSc

¹Institute of Information Systems Science, Department of Management and Entrepreneurship, Turku School of Economics, Turku, Finland

²Finland Futures Research Center, Turku School of Economics, University of Turku, Turku, Finland

Corresponding Author:

Reima Suomi, DSc

Institute of Information Systems Science

Department of Management and Entrepreneurship

Turku School of Economics

University of Turku

Turku, 20014

Finland

Phone: 358 50 410 2282

Fax: 358 2 333 5040

Email: reima.suomi@utu.fi

Abstract

Background: Social media services can help empower people to take greater responsibility for their health. For example, virtual worlds are media-rich environments that have many technically advantageous characteristics that can be used for Health 2.0 purposes. Second Life has been used to build environments where people can obtain information and interact with other users for peer support and advice from health care professionals.

Objective: The intent of the study was to find out whether Second Life is a working and functional platform supporting the empowerment of people in health-related issues.

Methods: We conducted a review of the current health-related activity in Second Life, coupled with an extensive series of observations and interactions with the respective resources inside Second Life.

Results: A total of 24 operative health resources were found in Second Life, indicating that health-related activity is rather limited in Second Life, though at first glance it appears to contain very rich health-related content. The other main shortcomings of Second Life relate to a lack of activity, a low number of resource users, problems with Second Life's search features, and the difficulty of finding trustworthy information.

Conclusions: For the average user, Second Life offers very little unique value compared to other online health resources.

(*J Med Internet Res* 2014;16(10):e229) doi:[10.2196/jmir.3177](https://doi.org/10.2196/jmir.3177)

KEYWORDS

social media; Second Life; virtual worlds; health; health promotion

Introduction

Importance of the Topic

Health 2.0 and Medicine 2.0 have become topical concepts in fields such as medicine, behavioral science, nursing studies, and information systems, to name a few [1-3]. Health 2.0 refers to “the use of a specific set of Web tools (blogs, podcasts, tagging, search, wikis, etc) by actors in health care, including doctors, patients, and scientists, using principles of open source and the generation of content by users, and the power of networks in order to personalize health care, collaborate, and promote health education”[3]. These concepts address the role of the individual in taking greater responsibility for their health

by utilizing technologies such as social media [4], defined as “a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, which allows the creation and exchange of user-generated content” [5].

For example, virtual worlds, defined as persistent computer-mediated 3-D environments where the users are represented as avatars [6], are a subset of social media [5] that have been used to build environments in which people can obtain and exchange health-related information. By offering a wide range of audio-visual stimuli in the media-rich, multi-user environment, they are able to create an immersive user experience and simulations of reality.

The characteristics of virtual worlds—the possibility to present oneself anonymously as an avatar, experience things, and interact with others in a virtual environment—can empower people with motoric or sensory disabilities and chronic illnesses that can hamper movement in the physical environment and face-to-face social interactions. For example, the virtual world platform Second Life hosts a wide range of health-related activities from virtual operating theatres for the purpose of education to places where people can look for information and receive support. Consequently, virtual worlds can help to empower individuals to take greater responsibility for their personal health and well-being [7].

Interestingly, despite their advantageous technical characteristics, only a little research has explored virtual worlds from the empowerment perspective. In more general terms, research about social media in health care is still in its infancy and lags far behind practice [8], although Beard et al [9] did survey health-related activities in Second Life when there was considerable hype around virtual worlds and Second Life.

Against this backdrop, the aim of this paper is twofold: first to conduct a follow-up study to Beard et al [9] to gain updated insight on the health-related activity in Second Life at a time when the hype around virtual worlds has largely vanished. Second, we focus on investigating how Second Life has been used for patient/citizen empowerment purposes. To this end, we have conducted an extensive set of observations in Second Life.

The remainder of the paper is organized as follows: after the introduction, we discuss patient empowerment and prior applications of Second Life in health care. In the third section, we present our empirical research and the results. In the fourth section, we discuss the implications of the results, present the limitations of the study, and suggest areas for further inquiry.

Background

Patient Empowerment With Social Media

People use social media in health care mainly for the purposes of sharing health care information and experiences with others, making effective decisions about health, and realizing the self-management of personal well-being. In the literature, the benefits of social media for the patient have been found to be: seeking emotional help or support, obtaining health information, exchanging information and experiences, seeking guidance online about health information and services, developing a different relationship with their doctors, health education and learning, and the self-management of health [1,2,10-14]. Social media has also been found to be helpful for different patient groups, such as teens with special health care needs as well as people with HIV, cancer, and chronic diabetes [13-15].

Social media can enable patients to become active and responsible partners in their health care activities [16]. As many diseases (eg, obesity, depression) are socially constructed, a key to their healing is also in social activity [17]. Thus, social media has been argued to have the potential to result in empowerment outcomes for patients wishing to improve communication with other patients and health professionals [18]. According to Eysenbach [4], patients using social media are empowered to contribute to and utilize the collective wisdom and experience of others to support their health care decisions and improve their health care.

Second Life and Health Care

Beard et al [9] surveyed health-related activities within Second Life in 2008. They discovered 68 health-related services in Second Life and concluded that the most common types of health-related resources in Second Life were those whose principle aim was to educate or offer information to patients or to increase awareness about health issues.

Yellowlees and Cook [19] investigated how Second Life can be used to increase awareness of severe mental illness. They constructed a visual environment to simulate the auditory and visual hallucinations of two patients with schizophrenia. A total of 863 survey respondents journeyed through the environment. Of those respondents, 440 (76%) thought the environment improved their understanding of auditory hallucinations, 69% thought it improved their understanding of visual hallucinations, and 82% said they would recommend the environment to a friend [19].

Boulos et al [1] evaluated the potential of 3-D environments in general in medical and health education and focused specifically on two resources in Second Life: Second Life Medical and Consumer Health Libraries and VNEC (Virtual Neurological Education Centre). They concluded (p. 242) that “3-D virtual worlds offer great potential to creative medical and health educators and librarians, but more research is needed into their use in medical and health education”.

The literature has also examined how Second Life could be used in treating patients. Gorini, Gaggioli, Vigna, and Riva [20] discussed the use of Second Life for psychotherapy. They found that compared to conventional telehealth applications, such as email, chat, and videoconferencing, the interaction between real and 3-D virtual worlds may convey greater feelings of presence between the therapists. In addition, Second Life was able to better facilitate the clinical communication process, positively influence group processes and cohesiveness in group-based therapies, and foster higher levels of interpersonal trust between patients (p. 21). [Figure 1](#) is a screenshot of the reception desk of Complete Care Medical Center Hospital.

Figure 1. Reception desk of Complete Care Medical Center Hospital.

Methods

Analysis of Current Health Activity in Second Life

Data Collection Phase

We collected the empirical data from health-related resources in Second Life. We first used the search features of Second Life as well as standard Internet searches to identify the health-related resources in Second Life. In the second stage of the empirical research, we conducted an extensive set of participant observation by accessing the resources we identified with our avatars. The data collection was initiated in late 2012 and concluded in June 2013.

We started our analysis by using the search features of Second Life (Second Life Destination Guide) with “health” as a search term. In addition, we did a similar search with other relevant keywords such as “hospital”, “medicine”, and “medical” to ensure we did not miss relevant results. Interestingly, we also found that the internal search features of Second Life returned slightly different numbers of results for analogous searches. Using the search features of Second Life, a total of 294 resources were found with the keyword “health”. Interestingly, in the Destination Guide of Second Life, using “health” as a search term did not return any results. In addition, we used standard Internet searches to obtain additional information about the resources we identified. This turned out to be very useful in evaluating the resources and further validating our findings.

We found that SL Healthy is a key resource within Second Life for health issues. The site itself describes itself as follows: “SL Healthy gathers information about consumer health locations and groups in Second Life, with general health education resources as well”. According to SL Healthy, there were 292 health resources in Second Life at the time of the data collection.

In addition, SL Science Center Group keeps track of information related to science and technology. SL Science Center Group list enumerates 11 resources with the title “clinic” and 2 resources with the title “health” (the list can be found with a Google search, not from Second Life). There is also a site nominating the top 10 Virtual Medical Sites in Second Life, although it was last updated in 2007.

To identify more relevant resources, we followed links, such as links to other Second Life locations, when applicable. For example, a good way to find new health-related activities was to follow other groups that the founders of health-related groups had belonged to. This approach is analogous to snowball sampling used in conducting a systematic literature review [21-23].

An overview of all health-related resources found in Second Life is presented in [Multimedia Appendix 1](#).

Data Cleaning Phase

The list of results from the Second Life search included parcels and groups. Parcels represent areas of virtual land in Second Life. They have required a financial input from their owners and their traffic is measured. Groups can be put up free of

charge, for example, around a topic of common interest and their activity level is impossible to follow without joining the groups.

The first cleaning task was to eliminate duplicates from the lists. This was difficult because the same resources sometimes appeared with slightly different names in different lists, and some resources formed nested structures.

After that, we excluded land parcels with a traffic rate of 20 or less based on Second Life statistics. This was due to the fact that these resources, although appearing in the search results, were however more or less “dead” with no customers. See [Figure 2](#) for a screenshot of the entrance to Global Health.

After having established the final list of relevant resources, four researchers entered Second Life and independently accessed

the listed resources and tabulated their findings. To systematize the analysis of the resources, we first adopted the classification of eHealth services by MacDonald, Case, and Mertzger [24] (see [Multimedia Appendix 2](#)). It soon became apparent that not all the components of the classification were applicable to our research context. For example, most of the groups were targeted to people who were both wealthy and ill, and material for prevention and cure was presented interchangeably. Thus, the researchers made notes about their observations and reflections when accessing the resources and saved material such as links and images in the project repository.

Thereafter, a final review and consolidation of the results was conducted. In the final stage, only 24 health-related active resources were identified and included for a further in-depth analysis as presented below.

Figure 2. Entrance of Global Health.



Results

Much of the activity in Second Life happens in different groups. [Table 1](#) presents their typical themes. Typically these resources were designed to raise awareness and provide information on a specific topic, such as pregnancy, alcoholism, or Asperger’s syndrome. In addition, many resources offered some form of peer support.

Resources promoting sexual health appeared in the results. However, many of the sexual health-related resources were put up to promote the rights of people with different sexual orientations and create venues for social interaction between

the members of these groups. Hence, deciding whether to include the theme of sexual health or not, and evaluating the degree of health orientation of this resource as well as others like it, was a challenging task.

Furthermore, there were also a handful of hospitals in Second Life. Their resources were often built to present new hospital concepts or to educate medical and nursing students on how to conduct medical procedures. Interestingly, it is rare for real life hospitals and private health companies to have a Second Life presence. This observation is in line with the findings by von Krogh, Jäger, and Barnetta [25] on the presence of Fortune 500 companies in Second Life.

A small group of resources concentrated on selling health-related items. These can be broadly divided into two very different categories: sellers of professional medical equipment such as instruments, clothing, and medical devices, and sellers offering different kinds of general products, such as scents, which they claimed contained therapeutic components. Interestingly, we did not observe any virtual pharmacy. We also observed resources for collecting money for charitable purposes, such as a resource offering residences, with the revenues being directed to medical research.

We also listed resources built around a specific illness or medical condition. [Table 2](#) offers an overview of those resources.

We found that the most active resources were networked in several ways. For example, a single institution might have set up several brands or “entrance points” to their services in real life and on the Internet. Additionally, many of the resources seemed to alter their names quite often, which caused additional challenges in the analysis.

Table 1. Examples of health-related groups in Second Life.

Type of resource	Description
Maternity	Information about pregnancy and maternity
Alcohol	Peer support and information for people suffering from alcohol abuse
Disabilities	Information and peer support for people with disabilities
Donations	Collecting money to support health-related activities and organizations

Table 2. Examples of illnesses and medical issues present in Second Life.

Illness/medical issue	Description
Cancer	Information about cancer treatments for patients; peer support for their families
Autism	General information; several groups
Dementia	General information
Diabetes	General information; peer support; information for parents of children with Type 1 diabetes
Mental health	Information about various mental health problems, eg, bipolar disorder, anxiety, and depression as well as the simulation of schizophrenia
Physical disabilities	Information, particularly for people using wheelchairs

Discussion

Comparison With Prior Research

In their study, Beard et al [9] discovered 68 health-related services in Second Life. The years between 2007 and 2008 are considered to be peak years of Second Life activity, since then the amount of users of health-related Second Life media have declined. Our study found just 24 active resources, confirming this trend.

Today, health-related activity in Second Life concentrates on issues such as accessibility, mental health, societal health, and peer-to-peer support. This matches current developments and needs in mainstream medicine, such as the Citizen-Centric Care paradigm [26-28]. In this line of thinking, health is essentially co-produced by individuals and their communities.

Universities seem to be the backbone of health and medical resources within Second Life because the platform is mainly used for teaching purposes. For example, the virtual hospital of Imperial College London offers a variety of simulated patient experiences for teaching students sequences of activities from the arrival of a patient.

One of the most interesting findings from our observations, when visiting the resources, was the lack of social interaction and the absence of other users. As Second Life is a very rich communication environment for facilitating the transmission of various social cues, such as facial expressions, gestures, and even voice, it offers many options for social interaction. From this perspective, it was surprising how little interaction we encountered during our observations. Most of the resources we visited were empty, except the welcoming robot greeting everyone entering the area. These observations do not mean that the resources are empty all the time; however, a lack of other users significantly reduces opportunities for peer support.

Key Findings

The main finding of this research was that virtual worlds such as Second Life can help people with illnesses and disabilities explore different things and interact with other people. In this sense, virtual worlds have the potential to help make societies more democratic by empowering people who have difficulties in real life. Furthermore, sensor-related disabilities can be compensated for by smart computer interfaces. Accordingly, a considerable part of the health-related activity in Second Life revolves around people with disabilities.

An interesting observation was that many of the health-related resources in Second Life were shut down or the level of activity appeared to be rather low. The technical capabilities of Second Life for offering users media-rich, immersive experiences in interaction with other users were not fully utilized. In most instances, we did not meet any other users. This indicates that

either Second Life or the maintainers of the health-related resources inside the virtual world have not been very successful in promoting the sustained usage of their services. This is well in line with the findings of Miller and Tucker [29], who found that much of social media traffic is generated by the social media-initiating firms themselves, not by real customers. If we consider that sustained usage, also referred to as continuance, has been viewed as a critical success factor for online services [30], the challenge of attracting users who can sustain services poses a major threat to Second Life.

Second Life has existed for only 10 years. The track records of most health care resources in Second Life give a picture of a haphazard entrance to the platform—one that has little emphasis on building long-term relationships with users and promoting continuance. An example of the lack of strategic orientation is that many resources have changed their names, which is obviously not an ideal approach when building a brand and awareness of it. Our research confirms the findings of Korda and Itani [31] that social media seldom achieve their desired outcomes in the health care field. A related challenge is contacting customers who have avatars that do not necessarily have any connection to their real-life identities. Building long-term relationships generally requires that the parties exchange information about themselves to develop trust [32]. Social media should be giving patients opportunities to assess the quality of health care offerings [33], but contradictory Second Life sites seem not to care much about their quality.

Our results portray a picture of Second Life as a dying environment full of deserted places and “ghost towns”. In most health-related resources, most or even all activities date back to between 2006 and 2008—more recent entries were hard to find. It must be noted that keeping time stamps is not one of the strengths of Second Life, which leads to the observation: what is a society that does not maintain its history or a track of time?

Limitations and Further Research

Our study has a number of limitations. First of all, as the environment constantly changes, our study is able to offer only a snapshot of a limited period of time. Second, determining the boundaries of the research area was problematic. For instance, does a discussion group on the beneficial effects of using horses to improve mental health fall within the research area? Much of the health-related activity in Second Life takes place in discussions between users, but these are not accessible without joining the (numerous) groups, so this method of validating the health-orientation of the resources was left unused in this research. In this study, we also focused only on the structures that have been built within the environment. Thus, a promising area for further research would be to conduct an ethnographic approach and observe active users, their communication patterns, and social structures in a longitudinal study. Third, in this study we focused solely on Second Life. However, there are a number of other virtual worlds that include health-related resources or facilitate health-related activities. For example, a leading social virtual world for teens, Habbo Hotel [34] has facilitated nurse

receptions to offer its young users an opportunity to ask health-related questions from a trained professional. Hence, we suggest future research with broader contextual coverage.

Conclusions

Based on the findings of this study, we argue that Second Life should be viewed as a stage in the evolution of virtual worlds. Since its inception, Second Life has educated large numbers of people about how to behave in the virtual environment and how to build a virtual presence. Today, its once innovative virtual world components, such as avatars and 3-D graphics, can be relatively easily implemented in any other online services.

Hence, we suggest that service developers analyze the shortcomings of Second Life in order to build better environments. For example, finding neutral statistics about Second Life was considerably challenging. Second, information resources such as SL Healthy were found to be rather untrustworthy as many of the listed resources had very little to do with health and many were inactive. Furthermore, there was very little transparency on how the list was constructed. Instead, developers of virtual environments should ensure that reliable, neutral, and up-to-date information is easily available. This would encourage organizations and businesses to invest in creating a presence on such platforms.

For developers of eHealth services, the technical components of virtual worlds such as Second Life enable the building of aesthetically impressive environments that provide rich communication possibilities, which in turn can promote sustained usage and service loyalty [35]. Thus, the lessons learned from Second Life can be useful for developing eHealth initiatives that help individuals to take greater responsibility for their health, promote the building of communities, and offer and receive peer support.

At the same time, learning features, such as controlling the avatar and operating in the virtual environment, requires effort and relationship-specific investment; both are likely to prohibit virtual worlds from diffusing to all segments of the population. The information richness facilitated by virtual worlds may be beneficial in some instances, whereas for those interested in obtain basic information about an illness or other medical condition, a standard website is in most cases a more effective and accessible solution. Furthermore, for building communities and interacting with other people, social networking sites such as Facebook are more convenient and already adopted by a significant proportion of the population. To conclude, based on our experiences and reflections from Second Life, we advise organizations developing eHealth initiatives to rigorously evaluate their features and user interface from a user perspective. Obtaining the user perspective can help the developers of eHealth initiatives to determine when the actual value of the service could be increased by implementing the media-rich features of virtual worlds, or whether a leaner media [36] and simpler user interface, such as a mobile app would be more appropriate to meet the needs of users.

Acknowledgments

This research has been performed without any direct external financing to this purpose. Matti Mäntymäki has been enjoying a 3-year research grant from Finnish Academy and Reima Suomi one year research grant from the Finnish Cultural Fund during the writing of this article. The authors wish to thank Mr Joonas Hamunen for the support in data collection from Second Life.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of health and medical-related information resources in Second Life.

[[JPG File, 94KB - jmir_v16i10e229_app1.JPG](#)]

Multimedia Appendix 2

Data collected from the Second Life Groups.

[[JPG File, 48KB - jmir_v16i10e229_app2.JPG](#)]

References

1. Boulos MNK, Hetherington L, Wheeler S. Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. *Health Info Libr J* 2007 Dec;24(4):233-245. [doi: [10.1111/j.1471-1842.2007.00733.x](https://doi.org/10.1111/j.1471-1842.2007.00733.x)] [Medline: [18005298](#)]
2. Huang E, Dunbar CL. Connecting to patients via social media: a hype or a reality? *Journal of Medical Marketing: Device, Diagnostic and Pharmaceutical Marketing* 2013 Mar 06;13(1):14-23. [doi: [10.1177/1745790413477647](https://doi.org/10.1177/1745790413477647)]
3. Hughes B, Joshi I, Wareham J. Health 2.0 and Medicine 2.0: tensions and controversies in the field. *J Med Internet Res* 2008;10(3):e23 [FREE Full text] [doi: [10.2196/jmir.1056](https://doi.org/10.2196/jmir.1056)] [Medline: [18682374](#)]
4. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. *J Med Internet Res* 2008;10(3):e22 [FREE Full text] [doi: [10.2196/jmir.1030](https://doi.org/10.2196/jmir.1030)] [Medline: [18725354](#)]
5. Kaplan A, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons* 2010;53(1):59-68 [FREE Full text]
6. Mäntymäki M, Salo J. Teenagers in social virtual worlds: Continuous use and purchasing behavior in Habbo Hotel. *Computers in Human Behavior* 2011 Nov;27(6):2088-2097 [FREE Full text] [doi: [10.1016/j.chb.2011.06.003](https://doi.org/10.1016/j.chb.2011.06.003)]
7. Bos L, Marsh A, Carroll D, Gupta S, Rees M. Patient 2.0 empowerment. 2008 Presented at: International Conference on Semantic Web & Web Services; July 14-17, 2008; Las Vegas, Nevada p. 164-167 URL: <http://members.media-effect.be/P4F/images/20100408patientempowermenthealth20.pdf>
8. Schein R, Wilson K, Keelan JE. A Report for Peel Public Health. Region of Peel: Peel Public Health; 2010. Literature review on effectiveness of the use of social media URL: <http://www.peelregion.ca/health/resources/pdf/socialmedia.pdf> [accessed 2014-09-09] [WebCite Cache ID 6ST9S4N6M]
9. Beard L, Wilson K, Morra D, Keelan J. A survey of health-related activities on Second Life. *J Med Internet Res* 2009 May;11(2):e17 [FREE Full text] [doi: [10.2196/jmir.1192](https://doi.org/10.2196/jmir.1192)] [Medline: [19632971](#)]
10. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. *Computers in Human Behavior* 2008 Sep;24(5):1867-1883. [doi: [10.1016/j.chb.2008.02.004](https://doi.org/10.1016/j.chb.2008.02.004)]
11. Coffield RL, Joiner JE. Risky business: Treating tweeting the symptoms of social media. *AHLA Connections* 2010;10-14 [FREE Full text]
12. Denecke K, Nejd W. How valuable is medical social media data? Content analysis of the medical web. *Information Sciences* 2009 May 30;179(12):1870-1880. [doi: [10.1016/j.ins.2009.01.025](https://doi.org/10.1016/j.ins.2009.01.025)]
13. Chira P, Nugent L, Miller K, Park T, Donahue S, Soni A, et al. Living Profiles: design of a health media platform for teens with special healthcare needs. *J Biomed Inform* 2010 Oct;43(5 Suppl):S9-12 [FREE Full text] [doi: [10.1016/j.jbi.2010.05.008](https://doi.org/10.1016/j.jbi.2010.05.008)] [Medline: [20937487](#)]
14. Newman MV, Lauterbach D, Munson SA, Resnick P, Morris ME. It's not that I don't have problems, I'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*. 2011 Presented at: ACM 2011; March 19-23, 2011; Hangzhou, China URL: http://delivery.acm.org/10.1145/1960000/1958876/p341-newman.pdf?ip=130.232.71.203&id=1958876&acc=ACTIVE%20SERVICE&key=74A0E95D84AAE420.469889998DA62376.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=566086938&CFTOKEN=30684076&acm=1411019630_5632e7cb8822a41b65945e0b5e6d9cdc [doi: [10.1145/1958824.1958876](https://doi.org/10.1145/1958824.1958876)]

15. Mo PK, Coulson NS. Empowering processes in online support groups among people living with HIV/AIDS: A comparative analysis of 'lurkers' and 'posters'. *Computers in Human Behavior* 2010 Sep;26(5):1183-1193. [doi: [10.1016/j.chb.2010.03.028](https://doi.org/10.1016/j.chb.2010.03.028)]
16. Lewis D. The multidimensional role of social media in healthcare. *Interactions* 2011 Jul 01;18(4):17. [doi: [10.1145/1978822.1978827](https://doi.org/10.1145/1978822.1978827)]
17. Coiera E. Social networks, social media, and social diseases. *BMJ* 2013 May 22;346:f3007-f3007. [doi: [10.1136/bmj.f3007](https://doi.org/10.1136/bmj.f3007)]
18. Bartlett YK, Coulson NS. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. *Patient Educ Couns* 2011 Apr;83(1):113-119. [doi: [10.1016/j.pec.2010.05.029](https://doi.org/10.1016/j.pec.2010.05.029)] [Medline: [20599338](https://pubmed.ncbi.nlm.nih.gov/20599338/)]
19. Yellowlees P, Cook J. Education about hallucinations using an internet virtual reality system: a qualitative survey. *Acad Psychiatry* 2006;30(6):534-539. [doi: [10.1176/appi.ap.30.6.534](https://doi.org/10.1176/appi.ap.30.6.534)] [Medline: [17139026](https://pubmed.ncbi.nlm.nih.gov/17139026/)]
20. 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](https://doi.org/10.2196/jmir.1029)] [Medline: [18678557](https://pubmed.ncbi.nlm.nih.gov/18678557/)]
21. Biernacki P, Waldorf D. Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods Research* 1981;10(2):141-163. [doi: [10.1177/004912418101000205](https://doi.org/10.1177/004912418101000205)]
22. Goodman LA. Snowball Sampling. *Ann Math Statist* 1961 Mar;32(1):148-170. [doi: [10.1214/aoms/1177705148](https://doi.org/10.1214/aoms/1177705148)]
23. Atkinson R, Flint J. Accessing hidden and hard-to-reach populations: Snowball research strategies. *Social Research Update* 2001(33):1-4.
24. MacDonald K, Case J, Mertzger J. E-Encounters. California, USA: California HealthCare Foundation; 2001.
25. von Krogh G, Jäger P, Barnetta A. Fortune 500 companies in Second Life - Activities, their success measurement and the satisfaction level of their projects. Zürich, Switzerland: Technische Hochschule Zürich; 2009. URL: http://www.smi.ethz.ch/education/thesis/Barnetta_SecondLife.pdf [accessed 2014-09-22] [WebCite Cache ID [6SmhX5fm6](https://www.webcitation.org/6SmhX5fm6)]
26. Woolf SH, Dekker MM, Byrne FR, Miller WD. Citizen-centered health promotion: building collaborations to facilitate healthy living. *Am J Prev Med* 2011 Jan;40(1 Suppl 1):S38-S47. [doi: [10.1016/j.amepre.2010.09.025](https://doi.org/10.1016/j.amepre.2010.09.025)] [Medline: [21146777](https://pubmed.ncbi.nlm.nih.gov/21146777/)]
27. Bertot JC, Jaeger PT, McClure CR. Citizen-centered e-government services: benefits, costs, and research needs. : Digital Government Society of North America; 2008 Nov Presented at: 2008 International Conference on Digital Government Research; May 18-21; Montreal, Canada URL: <http://dl.acm.org/citation.cfm?id=1367858>
28. Cooper TL, Bryet TA, Meek JW. Citizen-Centered Collaborative Public Management. *Public Administration Review* 2006 Dec;66(December):76-88.
29. Miller AR, Tucker C. Active social media management: the case of health care. *Information Systems Research* 2013 Mar;24(1):52-70. [doi: [10.1287/isre.1120.0466](https://doi.org/10.1287/isre.1120.0466)]
30. Bhattacharjee A. Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly* 2001;25(3):351-370. [doi: [10.2307/3250921](https://doi.org/10.2307/3250921)]
31. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. *Health Promot Pract* 2013 Jan;14(1):15-23. [doi: [10.1177/1524839911405850](https://doi.org/10.1177/1524839911405850)] [Medline: [21558472](https://pubmed.ncbi.nlm.nih.gov/21558472/)]
32. Morgan RM, Hunt SD. The commitment-trust theory of relationship marketing. *Journal of Marketing* 1994 Jul;58(3):20 [FREE Full text] [doi: [10.2307/1252308](https://doi.org/10.2307/1252308)]
33. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. *BMJ Qual Saf* 2013 Mar;22(3):251-255. [doi: [10.1136/bmjqs-2012-001527](https://doi.org/10.1136/bmjqs-2012-001527)] [Medline: [23349387](https://pubmed.ncbi.nlm.nih.gov/23349387/)]
34. Mäntymäki M, Riemer K. Digital natives in social virtual worlds: A multi-method study of gratifications and social influences in Habbo Hotel. *International Journal of Information Management* 2014 Apr;34(2):210-220. [doi: [10.1016/j.ijinfomgt.2013.12.010](https://doi.org/10.1016/j.ijinfomgt.2013.12.010)]
35. Cyr D, Hassanein K, Head M, Ivanov A. The role of social presence in establishing loyalty in e-Service environments. *Interacting with Computers* 2007 Jan;19(1):43-56. [doi: [10.1016/j.intcom.2006.07.010](https://doi.org/10.1016/j.intcom.2006.07.010)]
36. Daft RL, Lengel RH, Treviño LK. Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly* 1987;11(3):355-366 [FREE Full text]

Edited by G Eysenbach; submitted 11.12.13; peer-reviewed by R Kini, M Al-Hariri; comments to author 28.04.14; accepted 12.08.14; published 13.10.14.

Please cite as:

Suomi R, Mäntymäki M, Söderlund S

Promoting Health in Virtual Worlds: Lessons From Second Life

J Med Internet Res 2014;16(10):e229

URL: <http://www.jmir.org/2014/10/e229/>

doi: [10.2196/jmir.3177](https://doi.org/10.2196/jmir.3177)

PMID: [25313009](https://pubmed.ncbi.nlm.nih.gov/25313009/)

©Reima Suomi, Matti Mäntymäki, Sari Söderlund. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Automatic Evidence Retrieval for Systematic Reviews

Miew Keen Choong¹, PhD; Filippo Galgani¹, PhD; Adam G Dunn¹, PhD; Guy Tsafnat¹, PhD

Centre for Health Informatics, Australian Institute of Health Innovation, University of New South Wales, Kensington NSW, Australia

Corresponding Author:

Guy Tsafnat, PhD

Centre for Health Informatics

Australian Institute of Health Innovation

University of New South Wales

AGSM Building, Level 1

G27, Gate 11 Botany Street

Kensington NSW, 2052

Australia

Phone: 61 293858697

Fax: 61 293858692

Email: guyt@unsw.edu.au

Abstract

Background: Snowballing involves recursively pursuing relevant references cited in the retrieved literature and adding them to the search results. Snowballing is an alternative approach to discover additional evidence that was not retrieved through conventional search. Snowballing's effectiveness makes it best practice in systematic reviews despite being time-consuming and tedious.

Objective: Our goal was to evaluate an automatic method for citation snowballing's capacity to identify and retrieve the full text and/or abstracts of cited articles.

Methods: Using 20 review articles that contained 949 citations to journal or conference articles, we manually searched Microsoft Academic Search (MAS) and identified 78.0% (740/949) of the cited articles that were present in the database. We compared the performance of the automatic citation snowballing method against the results of this manual search, measuring precision, recall, and F₁ score.

Results: The automatic method was able to correctly identify 633 (as proportion of included citations: recall=66.7%, F₁ score=79.3%; as proportion of citations in MAS: recall=85.5%, F₁ score=91.2%) of citations with high precision (97.7%), and retrieved the full text or abstract for 490 (recall=82.9%, precision=92.1%, F₁ score=87.3%) of the 633 correctly retrieved citations.

Conclusions: The proposed method for automatic citation snowballing is accurate and is capable of obtaining the full texts or abstracts for a substantial proportion of the scholarly citations in review articles. By automating the process of citation snowballing, it may be possible to reduce the time and effort of common evidence surveillance tasks such as keeping trial registries up to date and conducting systematic reviews.

(*J Med Internet Res* 2014;16(10):e223) doi:[10.2196/jmir.3369](https://doi.org/10.2196/jmir.3369)

KEYWORDS

evidence-based medicine; medical informatics; information storage and retrieval

Introduction

Evidence retrieval tasks such as literature reviews and decision support, where recall of all relevant evidence is required, cannot rely on search technology alone due to limitations of keyword searching [1]. This has led to the development of secondary search methods such as citation tracking, called snowballing [2], and citation pearl growing [3,4].

Snowballing involves recursively pursuing relevant references cited in already-retrieved literature and adding them to the search results. Thus, snowballing is not limited to citation information found in bibliographical databases. The technical challenges of snowballing include obtaining the full text of retrieved citations, recognizing citation strings in the text, and retrieving new citations from citation strings. These make snowballing both tedious and time consuming.

Unlike keyword searching, snowballing does not require specific search terms [5], which are variable and inconsistent. Rather, it can be thought of as relying on the accumulation of multiple searches from different publishing authors [6]. Snowballing is a complementary approach to search for discovering additional evidence. Demonstrably, snowballing improves retrieval—for example, case studies using search and forward citation tracking on depression and coronary heart disease have been shown to identify more eligible articles than search alone [6]. A review on checking reference lists to find additional studies for systematic reviews found that citation tracking increased the yield of search results by 2.5-43% [7]. Snowballing is considered best practice and is frequently recommended when conducting systematic reviews [2]. With the rapid increase in the rate of publication, automation of snowballing would significantly reduce the time dedicated to literature search.

We tested an approach to automatic snowballing that uses citation extraction algorithms and scientific search engines to follow the steps of snowballing: (1) extract citation strings from documents, (2) find the citations, (3) fetch the full text of citations, and (4) repeat the process to recursively retrieve more citations. The goal of this study is to test the feasibility of automating each of the subtasks of snowballing.

Methods

Algorithm

With an initial set of at least one paper, portable document format (PDF) and hypertext markup language (HTML) documents are converted to plain text. A modified version of ParsCit [8] (a free and open source reference strings segmentation package) identifies reference sections and segments individual reference strings. We modified the heuristics in ParsCit and changed the restriction by allowing the reference label to be found from the middle (50%) to the end of the text. Each reference is converted to a search engine query by removing short words, numbers, and punctuation. The query results returned from the search engine contain citation information (eg, authors, titles, journal, year, digital object identifiers [DOI]) and often one or more links (uniform resource locator [URL]) to full text. We extracted and followed links to the full text. The source code is available in [Multimedia Appendix 1](#) and [9].

Data

In the evaluation, we used citations from a set of published English language reviews about neuraminidase inhibitors. The dataset consisted of 152 systematic and non-systematic review articles. We randomly selected a subset of 20 review articles that contained 1057 citations. We excluded references to websites, books, book chapters, newspaper articles, and grey literature, leaving 949 included citations. The properties of the 20 review articles are provided in [Multimedia Appendix 2](#).

Study Design

We evaluated our algorithm using the proportion of extracted references, the proportion of citations retrieved, and the proportion of abstracts and full texts downloaded. We checked extracted citations manually against the references in the paper. We considered a reference to be correctly extracted only if it contained the entire reference without loss of information. We did allow for minimal extra information, such as white space and citation number but not information that should have been part of another citation string, page footer or manuscript text. The accuracy of the retrieved citations and abstract/full text with the references from the systematic reviews were verified manually. Correctly retrieved articles were counted as true positives. Retrieved articles that are not the ones cited were counted as false positives.

We used Microsoft Academic Search (MAS) [10] ([Figure 1](#)) in the evaluation. MAS is a generalized scientific literature search engine that covers more than 48 million publications with weekly updates. A free application programming interface (API) is provided for non-commercial purposes after registration. Citations include bibliographic data as well as links to citing papers and links to multiple versions of the paper if more than one version exists, including, often, to full text. We used the MAS API to perform searches for each of the references extracted from the full text of the original paper. Other search engines (eg, Google Scholar [11]) can also be used in this step, subject to restrictions they impose. We chose MAS due to its size, “cited by” functionality, links to full text, and because it does not enforce active blocking to prevent automated access.

We manually searched for missed references to ascertain whether they were indeed indexed in MAS. Articles that were not retrieved but were found by manual search of MAS were counted as false negatives. We calculated precision, recall, and F_1 score using the standard formulae:

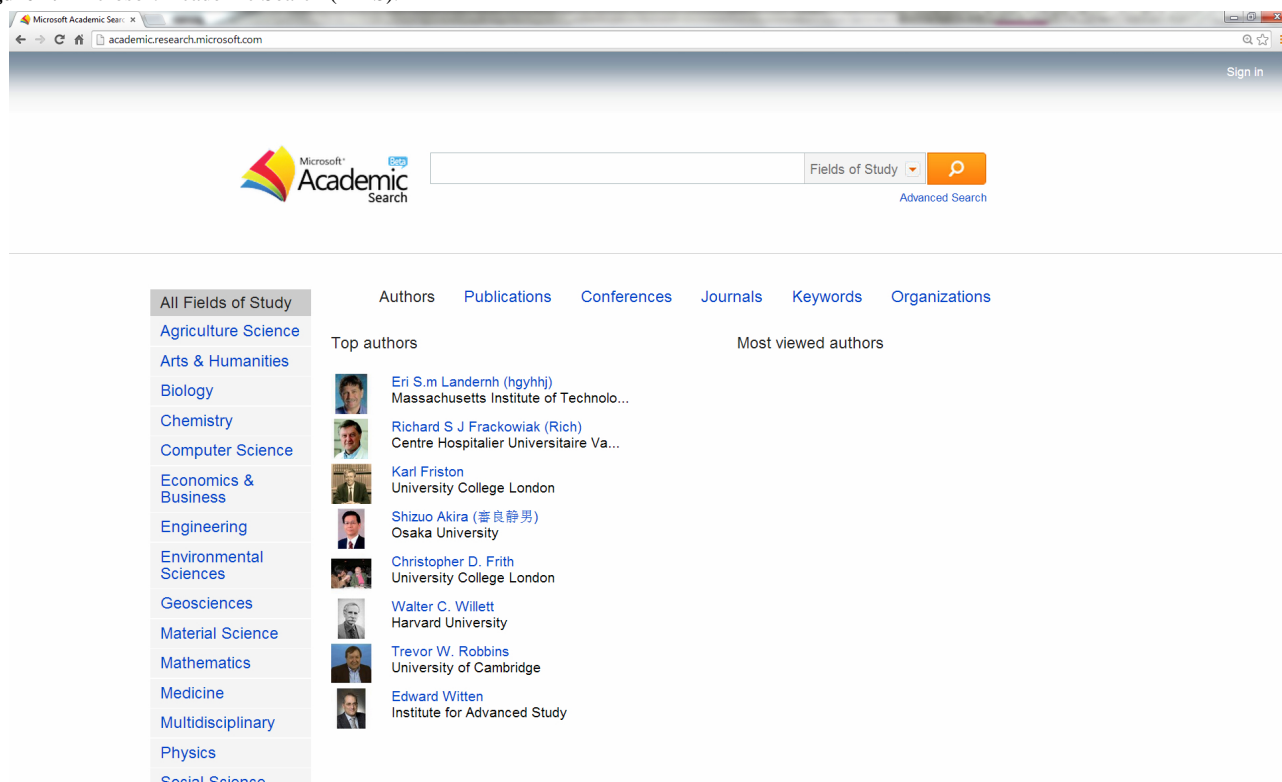
$$\text{Precision} = (\text{True positives}) / (\text{True positives} + \text{False positives})$$

$$\text{Recall} = (\text{True positives}) / (\text{True positives} + \text{False negatives})$$

$$F_1 \text{ score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$$

The precision, recall, and F_1 scores were computed for retrieval of citations, abstract (only abstracts or abstracts with full texts), and full text against all citations (1057 references), included citations (949 references), and included citations indexed in MAS (740 references).

All experiments were conducted on computers with Internet protocols (IP) allocated to the University of New South Wales. Journals that automatically recognize subscription by IP address and to which the University of New South Wales library is subscribed were thus granted access. No other subscription activation or authentication methods were used. However, since most abstracts are freely accessible, download of abstracts will not normally be affected by journal subscription.

Figure 1. Microsoft Academic Search (MAS).

Results

The summary of the evaluation is shown in [Figure 2](#). We successfully extracted 97.2% (922/949) of the included citations (96.5%, 1020/1057 citations) from the 20 reviews randomly selected. The precision, recall and F_1 scores for retrieval of citations, abstract (only abstracts or abstracts with full texts), and full text fetching of citations from the 20 randomly selected review articles are shown in [Table 1](#). The scores are computed using three categories: (1) all citations, (2) included citations, and (3) included citations indexed in MAS. The first category shows the probability to retrieve a given reference from a review paper. The second category gives the same probability but excludes works not likely to be retrievable such as books and grey literature. The third category excludes citations not currently indexed in MAS.

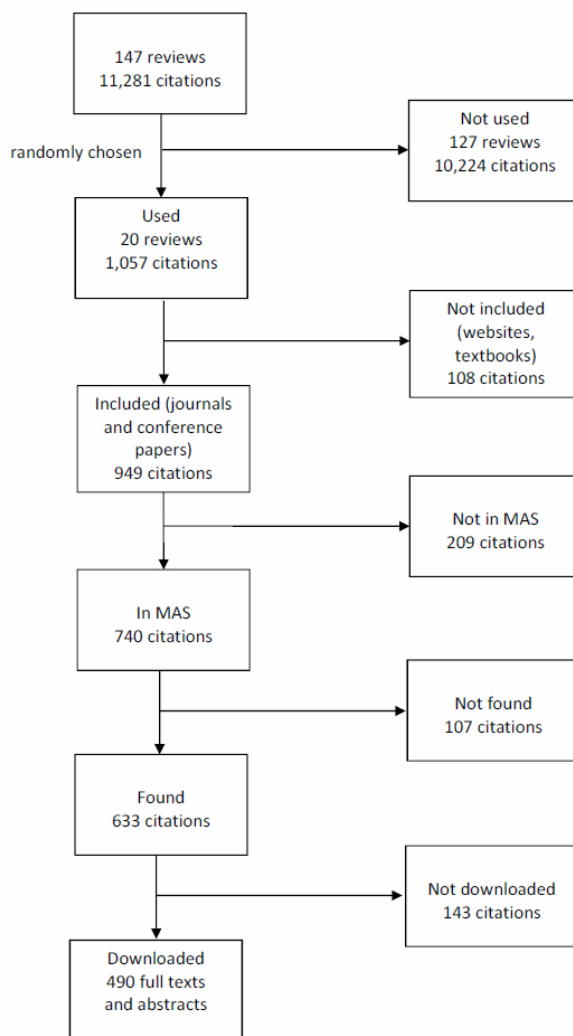
For the reference strings indexed in MAS, 66.2% (490/740) of abstracts were correctly downloaded either on their own or as part of the full text. These represent 51.6% of 949 included citations and 46.4% of all 1057 references included in the study.

Out of the 633 correctly identified citations, we retrieved the full text or abstract for 490 (recall=82.9%, precision=92.1%, F_1 score=87.3%). We examined the specific reasons why 143 (22.6%) of the articles were not downloaded automatically. MAS had incorrect links for 39 articles (6.2%), and no link to full text for 6 articles (0.9%); 56 links (8.8%) led to a page that uses JavaScript to dynamically create a link to the full text. For citations where only abstracts were downloaded (15 abstracts), full text documents were not downloaded due to journal subscription access.

Table 1. Results of citations, abstract, and full text retrieval (precision, recall, and F₁ score of database results fetch, and full text and abstract retrieval, comparing all reference strings, only included citations, and only included citations indexed in MAS).

	As proportion of all citations (n=1057)	As proportion of included citations (n=949)	As proportion of citations in MAS (n=740)
Citations retrieved			
Precision	0.977	0.977	0.977
Recall	0.600	0.667	0.855
F ₁ score	0.743	0.793	0.912
Abstracts fetched			
Precision	0.921	0.921	0.921
Recall	0.483	0.540	0.702
F ₁ score	0.634	0.681	0.797
Full text fetched			
Precision	0.919	0.919	0.919
Recall	0.475	0.533	0.696
F ₁ score	0.626	0.674	0.792

Figure 2. Summary of the evaluation results (from 20 reviews with 949 scholarly citations, MAS included 740 citations, 633 citations were found, and 490 full texts and abstracts were downloaded automatically).



Discussion

Principal Findings

Snowballing is tedious and resource demanding but has shown to improve retrieval. This evaluation shows that it is feasible to automatically perform snowballing using our method by extracting and downloading the citations. Systems designed to perform many of the systematic review tasks are already in use, in development, or in research [12,13]. This study represents a first effort to automate the snowballing tasks in a systematic review process. When integrated with a reliable automatic screening tool, automatic snowballing can have a compound effect and increase recall [12].

Automatic citation extraction is a difficult task [14], which causes the citation retrieval to be an even harder task. However, if unique identifier of citations (eg, DOI or PubMed identifier) is provided for each citation, this would greatly improve the reliability of citation extraction and retrieval.

Limitations

A limitation of this study is that full text fetching is tested on journal subscription by IP address and to which the University of New South Wales library is subscribed. While this means

that results may vary in other institutions, they also represent an exemplar that may guide expectations of results. With the growth of open source and other means of obtaining full text [15], the performance of our algorithm may improve.

In this evaluation, the algorithm was limited to MAS. This is a constraint of the testing system, not of the method. From the limited testing we have conducted, the algorithm performs equivalently on Google Scholar but computer-access restrictions prevented a robust comparison.

Some existing databases, such as Scopus [16] and Web of Science [17] (subscription fees apply for both), provide citation analysis and allow one to search both forward (references cited in an investigated text) and backward (papers citing an investigated text) and can thus aid manual snowballing. However, those citations are limited to papers indexed in the respective database. Our method automatically extracts citations directly from documents and can thus cross database boundaries.

Conclusions

Snowballing is automatable and can reduce the time and effort of evidence retrieval. It is possible to reliably extract reference lists from the text of scientific papers, find these citations in scientific search engines, and fetch the full text and/or abstract.

Acknowledgments

This work was supported by a National Health & Medical Research Council Centre for Research Excellence in eHealth Grant APP1032664.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Source code <http://www2.chi.unsw.edu.au/~miewkeen/ESuRFr.html>.

[ZIP File (Zip Archive), 1MB - [jmir_v16i10e223_app1.zip](#)]

Multimedia Appendix 2

Properties of the 20 review articles included in the study.

[PDF File (Adobe PDF File), 3KB - [jmir_v16i10e223_app2.pdf](#)]

References

1. Ceri S, Bozzon A, Brambilla M, Della Valle E, Fraternali P, Quarteroni S. The information retrieval process. In: Carey MJ, Ceri S, editors. *Web Information Retrieval*. Heidelberg: Springer; 2013:13-26.
2. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005 Nov 5;331(7524):1064-1065 [FREE Full text] [doi: [10.1136/bmj.38636.593461.68](https://doi.org/10.1136/bmj.38636.593461.68)] [Medline: [16230312](https://pubmed.ncbi.nlm.nih.gov/16230312/)]
3. Ramer SL. Site-ation pearl growing: methods and librarianship history and theory. *J Med Libr Assoc* 2005 Jul;93(3):397-400 [FREE Full text] [Medline: [16059431](https://pubmed.ncbi.nlm.nih.gov/16059431/)]
4. Search Strategies and Heuristics for Searching Bibliographic and Full Text Databases. URL: <https://dspace.creighton.edu/xmlui/bitstream/handle/10504/5903/Searching%20Databases%20-%20Strategies%20and%20Heuristics.pdf?sequence=1> [accessed 2014-08-12] [WebCite Cache ID 6RIPHoDrK]
5. Robinson KA, Dunn AG, Tsafnat G, Glasziou P. Citation networks of related trials are often disconnected: implications for bidirectional citation searches. *J Clin Epidemiol* 2014 Jul;67(7):793-799. [doi: [10.1016/j.jclinepi.2013.11.015](https://doi.org/10.1016/j.jclinepi.2013.11.015)] [Medline: [24725642](https://pubmed.ncbi.nlm.nih.gov/24725642/)]

6. Kuper H, Nicholson A, Hemingway H. Searching for observational studies: what does citation tracking add to PubMed? A case study in depression and coronary heart disease. *BMC Med Res Methodol* 2006;6:4 [FREE Full text] [doi: [10.1186/1471-2288-6-4](https://doi.org/10.1186/1471-2288-6-4)] [Medline: [16483366](#)]
7. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev* 2011(8):MR000026. [doi: [10.1002/14651858.MR000026.pub2](https://doi.org/10.1002/14651858.MR000026.pub2)] [Medline: [21833989](#)]
8. Councill IG, Giles CL, Kan MY. ParsCit: an open-source CRF reference string parsing package. In: Proceedings of the Sixth International Language Resources and Evaluation. 2008 Presented at: International Language Resources and Evaluation Conference; May 2008; Marrakesh, Morocco.
9. Evidence Discovery using Microsoft Academic Search (MAS) Sample Code version 0.1. URL: <http://www2.chi.unsw.edu.au/~miewkeen/ESuRFr.html> [accessed 2014-08-12] [WebCite Cache ID [6RILz49RD](#)]
10. Microsoft Academic Search. URL: <http://academic.research.microsoft.com/> [accessed 2014-07-14] [WebCite Cache ID [6R35GdAH5](#)]
11. Google Scholar. URL: <http://scholar.google.com.au/> [accessed 2014-07-14] [WebCite Cache ID [6R35D2LI2](#)]
12. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. *Syst Rev* 2014;3:74 [FREE Full text] [doi: [10.1186/2046-4053-3-74](https://doi.org/10.1186/2046-4053-3-74)] [Medline: [25005128](#)]
13. Wallace BC, Dahabreh IJ, Schmid CH, Lau J, Trkalinos TA. Modernizing the systematic review process to inform comparative effectiveness: tools and methods. *J Comp Eff Res* 2013 May;2(3):273-282. [doi: [10.2217/ce.13.17](https://doi.org/10.2217/ce.13.17)] [Medline: [24236626](#)]
14. Day MY, Tsai TH, Sung CL, Lee CW, Wu SH, Ong CS, et al. A knowledge-based approach to citation extraction. 2005 Presented at: IEEE International Conference on Information Reuse and Integration; August 15-17, 2005; Las Vegas, NV p. 50-55. [doi: [10.1109/IRI-05.2005.1506448](https://doi.org/10.1109/IRI-05.2005.1506448)]
15. Dunn AG, Coiera E, Mandl KD. Is Biblioleaks inevitable? *J Med Internet Res* 2014;16(4):e112 [FREE Full text] [doi: [10.2196/jmir.3331](https://doi.org/10.2196/jmir.3331)] [Medline: [24755534](#)]
16. Scopus. URL: <http://www.scopus.com/> [accessed 2014-07-14] [WebCite Cache ID [6R35JZXxj](#)]
17. Web of Science. URL: <http://thomsonreuters.com/thomson-reuters-web-of-science/> [accessed 2014-07-14] [WebCite Cache ID [6R35Kw0b4](#)]

Abbreviations

- API:** application programming interface
DOI: digital object identifier
IP: Internet protocol
MAS: Microsoft Academic Search

Edited by G Eysenbach; submitted 14.07.14; peer-reviewed by R Badgett, B Wallace; comments to author 07.08.14; revised version received 18.08.14; accepted 09.09.14; published 01.10.14.

Please cite as:

Choong MK, Galgani F, Dunn AG, Tsafnat G
Automatic Evidence Retrieval for Systematic Reviews
J Med Internet Res 2014;16(10):e223
URL: <http://www.jmir.org/2014/10/e223/>
doi: [10.2196/jmir.3369](https://doi.org/10.2196/jmir.3369)
PMID: [25274020](https://pubmed.ncbi.nlm.nih.gov/25274020/)

©Miew Keen Choong, Filippo Galgani, Adam G Dunn, Guy Tsafnat. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.10.2014. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>