
Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 15 (2013), Issue 8 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

- Validity of Web-Based Self-Reported Weight and Height: Results of the Nutrinet-Santé Study ([e152](#))
Camille Lassale, Sandrine Péneau, Mathilde Touvier, Chantal Julia, Pilar Galan, Serge Hercberg, Emmanuelle Kesse-Guyot. 5
- Determinants of Participation in a Web-Based Health Risk Assessment and Consequences for Health Promotion Programs ([e151](#))
Maurice Niessen, Eva Laan, Suzan Robroek, Marie-Louise Essink-Bot, Niels Peek, Roderik Kraaijenhagen, Coen Van Kalken, Alex Burdorf. 1 8
- Usage of a Generic Web-Based Self-Management Intervention for Breast Cancer Survivors: Substudy Analysis of the BREATH Trial ([e170](#))
Sanne van den Berg, Esmee Peters, J Kraaijeveld, Marieke Gielissen, Judith Prins. 31
- Pedometer-Based Internet-Mediated Intervention For Adults With Chronic Low Back Pain: Randomized Controlled Trial ([e181](#))
Sarah Krein, Reema Kadri, Maria Hughes, Eve Kerr, John Piette, Rob Holleman, Hyungjin Kim, Caroline Richardson. 42
- Use of a Web 2.0 Portal to Improve Education and Communication in Young Patients With Families: Randomized Controlled Trial ([e175](#))
Lena Hanberger, Johnny Ludvigsson, Sam Nordfeldt. 56
- Barriers and Facilitators for the Implementation of an Online Clinical Health Community in Addition to Usual Fertility Care: A Cross-Sectional Study ([e163](#))
Johanna Aarts, Marjan Faber, Anne den Boogert, Ben Cohlen, Paul van der Linden, Jan Kremer, Willianne Nelen. 70
- Understanding Determinants of Consumer Mobile Health Usage Intentions, Assimilation, and Channel Preferences ([e149](#))
Arun Rai, Liwei Chen, Jessica Pye, Aaron Baird. 84
- Tablet-Based Strength-Balance Training to Motivate and Improve Adherence to Exercise in Independently Living Older People: A Phase II Preclinical Exploratory Trial ([e159](#))
Patrícia Silveira, Rolf van de Langenberg, Eva van het Reve, Florian Daniel, Fabio Casati, Eling de Bruin. 104
- Standardized, App-Based Disinfection of iPads in a Clinical and Nonclinical Setting: Comparative Analysis ([e176](#))
Urs-Vito Albrecht, Ute von Jan, Ludwig Sedlacek, Stephanie Groos, Sebastian Suerbaum, Ralf-Peter Vonberg. 121

Efficacy of a Text Message-Based Smoking Cessation Intervention for Young People: A Cluster Randomized Controlled Trial (e171)	
Severin Haug, Michael Schaub, Vigeli Venzin, Christian Meyer, Ulrich John.	132
Enhancing Retention of an Internet-Based Cohort Study of Men Who Have Sex With Men (MSM) via Text Messaging: Randomized Controlled Trial (e194)	
Christine Khosropour, Brent Johnson, Alexandra Ricca, Patrick Sullivan.	146
Use of Twitter Among Local Health Departments: An Analysis of Information Sharing, Engagement, and Action (e177)	
Brad Neiger, Rosemary Thackeray, Scott Burton, Callie Thackeray, Jennifer Reese.	157
The Clinical Effectiveness of Web-Based Cognitive Behavioral Therapy With Face-to-Face Therapist Support for Depressed Primary Care Patients: Randomized Controlled Trial (e153)	
Ragnhild Høifødt, Kjersti Lillevoll, Kathleen Griffiths, Tom Wilsgaard, Martin Eisemann, Knut Waterloo, Nils Kolstrup.	174
Staging Dementia From Symptom Profiles on a Care Partner Website (e145)	
Kenneth Rockwood, Matthew Richard, Chris Leibman, Lisa Mucha, Arnold Mitnitski.	196
Internet-Based Early Intervention to Prevent Posttraumatic Stress Disorder in Injury Patients: Randomized Controlled Trial (e165)	
Joanne Mouthaan, Marit Sijbrandij, Giel-Jan de Vries, Johannes Reitsma, Rens van de Schoot, J Goslings, Jan Luitse, Fred Bakker, Berthold Gersons, Miranda Olf.	206
Participants, Usage, and Use Patterns of a Web-Based Intervention for the Prevention of Depression Within a Randomized Controlled Trial (e172)	
Saskia Kelders, Ernst Bohlmeijer, Julia Van Gemert-Pijnen.	219
Using Twitter to Examine Smoking Behavior and Perceptions of Emerging Tobacco Products (e174)	
Mark Myslín, Shu-Hong Zhu, Wendy Chapman, Mike Conway.	234
Patient and Public Views on Electronic Health Records and Their Uses in the United Kingdom: Cross-Sectional Survey (e160)	
Serena Luchenski, Julie Reed, Cicely Marston, Chrysanthi Papoutsis, Azeem Majeed, Derek Bell.	250
Limited Use of Price and Quality Advertising Among American Hospitals (e185)	
David Muhlestein, Chrisanne Wilks, Jason Richter.	263
The MONitoring Resynchronization dEVICES and CARdiac patiEnts (MORE-CARE) Randomized Controlled Trial: Phase 1 Results on Dynamics of Early Intervention With Remote Monitoring (e167)	
Giuseppe Boriani, Antoine Da Costa, Renato Ricci, Aurelio Quesada, Stefano Favale, Saverio Iacopino, Francesco Romeo, Arnaldo Risi, Lorenza Mangoni di S Stefano, Xavier Navarro, Mauro Biffi, Massimo Santini, Haran Burri, On Behalf Of The MORE-CARE Investigators.	273
Web-Based Versus Traditional Paper Questionnaires: A Mixed-Mode Survey With a Nordic Perspective (e173)	
Lena Hohwü, Heidi Lyshol, Mika Gissler, Stefan Jonsson, Max Petzold, Carsten Obel.	288
Using the Internet to Help With Diet, Weight, and Physical Activity: Results From the Health Information National Trends Survey (HINTS) (e148)	
Scout McCully, Brian Don, John Updegraff.	299
Consumers' Perceptions of Patient-Accessible Electronic Medical Records (e168)	
Christina Zarcadoolas, Wendy Vaughn, Sara Czaja, Joslyn Levy, Maxine Rockoff.	307

Can Social Semantic Web Techniques Foster Collaborative Curriculum Mapping In Medicine? (e169) Cord Spreckelsen, Sonja Finsterer, Jan Cremer, Hennig Schenkat.	324
Analysis of the Security and Privacy Requirements of Cloud-Based Electronic Health Records Systems (e186) Joel JPC Rodrigues, Isabel de la Torre, Gonzalo Fernández, Miguel López-Coronado.	348
Attrition Revisited: Adherence and Retention in a Web-Based Alcohol Trial (e162) Elizabeth Murray, Ian White, Mira Varagunam, Christine Godfrey, Zarnie Khadjesari, Jim McCambridge.	357
Retrieving Clinical Evidence: A Comparison of PubMed and Google Scholar for Quick Clinical Searches (e164) Salimah Shariff, Shayna Bejaimal, Jessica Sontrop, Arthur Iansavichus, R Haynes, Matthew Weir, Amit Garg.	368
A Method for the Design and Development of Medical or Health Care Information Websites to Optimize Search Engine Results Page Rankings on Google (e183) Suzanne Dunne, Niamh Cummins, Ailish Hannigan, Bill Shannon, Colum Dunne, Walter Cullen.	377
An Analysis of Online Evaluations on a Physician Rating Website: Evidence From a German Public Reporting Instrument (e157) Martin Emmert, Florian Meier.	385
Physician Choice Making and Characteristics Associated With Using Physician-Rating Websites: Cross-Sectional Study (e187) Martin Emmert, Florian Meier, Frank Pisch, Uwe Sander.	396
Crowdsourcing Black Market Prices For Prescription Opioids (e178) Nabarun Dasgupta, Clark Freifeld, John Brownstein, Christopher Menone, Hilary Surratt, Luke Poppish, Jody Green, Eric Lavonas, Richard Dart.	412
Effects of Mobile Augmented Reality Learning Compared to Textbook Learning on Medical Students: Randomized Controlled Pilot Study (e182) Urs-Vito Albrecht, Kristian Folta-Schoofs, Marianne Behrends, Ute von Jan.	423
 Review	
The Impact of Social Media on Medical Professionalism: A Systematic Qualitative Review of Challenges and Opportunities (e184) Fatemeh Gholami-Kordkheili, Verina Wild, Daniel Strech.	166
 Viewpoints	
Security Analysis and Improvements to the PsychoPass Method (e161) Bostjan Brumen, Marjan Heri ko, Ivan Rozman, Marko Hölbl.	341
The Emergent Discipline of Health Web Science (e166) Joanne Luciano, Grant Cumming, Mark Wilkinson, Eva Kahana.	436



Letters to the Editor

Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research (e179) John Cunningham.	445
Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research - Author's Reply (e180) Reid Hester, William Campbell, Kathryn Lenberg, Harold Delaney.	447

Original Paper

Validity of Web-Based Self-Reported Weight and Height: Results of the Nutrinet-Santé Study

Camille Lassale¹, MSc Public Health; Sandrine Péneau¹, PhD; Mathilde Touvier¹, PhD; Chantal Julia^{1,2}, MSc Public Health, MD; Pilar Galan¹, MD, PhD; Serge Hercberg^{1,2}, MD, PhD; Emmanuelle Kesse-Guyot¹, PhD

¹Université Paris 13, Sorbonne Paris Cité, UREN (Nutritional Epidemiology Research Unit), Inserm (U557), Inra (U1125), Cnam, Bobigny, France

²Public Health Department, Hôpital Avicenne, Bobigny, France

Corresponding Author:

Camille Lassale, MSc Public Health

Université Paris 13, Sorbonne Paris Cité, UREN (Nutritional Epidemiology Research Unit), Inserm (U557), Inra (U1125), Cnam
74 rue Marcel Cachin

Bobigny, 93017

France

Phone: 33 148388974

Fax: 33 148388931

Email: c.lassale@uren.smbh.univ-paris13.fr

Abstract

Background: With the growing scientific appeal of e-epidemiology, concerns arise regarding validity and reliability of Web-based self-reported data.

Objective: The objectives of the present study were to assess the validity of Web-based self-reported weight, height, and resulting body mass index (BMI) compared with standardized clinical measurements and to evaluate the concordance between Web-based self-reported anthropometrics and face-to-face declarations.

Methods: A total of 2513 participants of the NutriNet-Santé study in France completed a Web-based anthropometric questionnaire 3 days before a clinical examination (validation sample) of whom 815 participants also responded to a face-to-face anthropometric interview (concordance sample). Several indicators were computed to compare data: paired t test of the difference, intraclass correlation coefficient (ICC), and Bland–Altman limits of agreement for weight, height, and BMI as continuous variables; and kappa statistics and percent agreement for validity, sensitivity, and specificity of BMI categories (normal, overweight, obese).

Results: Compared with clinical data, validity was high with ICC ranging from 0.94 for height to 0.99 for weight. BMI classification was correct in 93% of cases; kappa was 0.89. Of 2513 participants, 23.5% were classified overweight (BMI \geq 25) with Web-based self-report vs 25.7% with measured data, leading to a sensitivity of 88% and a specificity of 99%. For obesity, 9.1% vs 10.7% were classified obese (BMI \geq 30), respectively, leading to sensitivity and specificity of 83% and 100%. However, the Web-based self-report exhibited slight underreporting of weight and overreporting of height leading to significant underreporting of BMI ($P<.05$) for both men and women: -0.32 kg/m² (SD 0.66) and -0.34 kg/m² (SD 1.67), respectively. Mean BMI underreporting was -0.16 , -0.36 , and -0.63 kg/m² in the normal, overweight, and obese categories, respectively. Almost perfect agreement (ie, concordance) was observed between Web-based and face-to-face report (ICC ranged from 0.96 to 1.00, classification agreement was 98.5%, and kappa 0.97).

Conclusions: Web-based self-reported weight and height data from the NutriNet-Santé study can be considered as valid enough to be used when studying associations of nutritional factors with anthropometrics and health outcomes. Although self-reported anthropometrics are inherently prone to biases, the magnitude of such biases can be considered comparable to face-to-face interview. Web-based self-reported data appear to be an accurate and useful tool to assess anthropometric data.

(*J Med Internet Res* 2013;15(8):e152) doi:[10.2196/jmir.2575](https://doi.org/10.2196/jmir.2575)

KEYWORDS

anthropometry; body weight; obesity; self-report; weights and measures; validation studies

Introduction

Overweight and obesity have reached pandemic proportions and it is considered as one of the major public health issues by the World Health Organization (WHO) [1-3]. Excess body weight is a major risk factor of various chronic conditions, such as hypertension, type 2 diabetes, cardiovascular diseases, and some cancers [4].

Body mass index (BMI), defined as weight (kg) divided by squared height (m^2), is highly correlated to excess fat mass. It is commonly used to classify overweight and obesity in adults: overweight excluding obesity (BMI 25-29 kg/m^2) and obesity (BMI ≥ 30 kg/m^2) [1]. In large-scale multicentric epidemiologic studies, self-reporting of weight and height is usually used because of substantial logistic and cost savings as compared with direct measures by trained technicians. In that context, self-reporting is actually the more effective and manageable way to collect anthropometric data in large samples up to tens of thousands of participants.

However, it is acknowledged that self-reported height and weight are biased proxies of the true measures. Indeed, bias between self-reported and measured anthropometrics has been widely described in the scientific literature, in many American and European studies [5-13]. Generally, weight is underreported whereas height is overreported, [5,12] leading to an underestimation of BMI and a misclassification in BMI categories, although errors vary according to sex, age, education, and socioeconomic characteristics [8,10,11,14,15]. Moreover, biases are likely differential with a relationship between magnitude of bias and measured BMI: underweight participants tend to overreport whereas overweight participants tend to underreport their weight [16]. This phenomenon is partly explained by social desirability, which can be further influenced by the method of data collection [5,7,17,18]. For example, evidence for social desirability bias was observed in the Canadian Community Health Survey, which studied the difference between face-to-face and telephone self-reported anthropometrics and showed that obesity prevalence in the face-to-face group was significantly higher than in the phone group (18% and 13%, respectively) [18]. This suggests a tendency to underreport weight to attempt to construct favorable images in the eyes of others, to get closer to a socially ideal weight when the interviewer cannot visually assess it [19]. In that context, it is of interest to assess whether Web-based self-report would lead to the same discrepancy with face-to-face compared to what is observed between telephone and face-to-face self-report.

A novel approach for large-scale epidemiologic studies lies in the use of Internet to administer Web-based questionnaires [20-25], which is recognized as the new promising field of e-epidemiology. A key advantage of a Web-based epidemiologic study is the substantial logistic and cost savings compared with traditional data collection (pencil and paper questionnaires, face-to-face interviews). Other features, such as data management improvement and simplification, flexibility, and recruitment of large samples, can be achieved with e-epidemiology.

In the NutriNet-Santé study, comparison of self-reported weight and height in a Web-based anthropometric questionnaire with the traditional paper form of the same questionnaire showed satisfying results, which were published elsewhere [26].

To date, only 1 study focused on assessing validity of Web-based self-reported weight compared with direct measure [27]. However, this study did not provide insight on the validity of Web-based self-reported height or BMI because height was not measured. To the best of our knowledge, the comparison between Web-based and face-to-face self-reported anthropometrics has never been published.

The objectives of the present study were to (1) assess the validity of Web-based self-reported weight, height, and resulting BMI compared with measured data in a subsample of the NutriNet-Santé study, and (2) evaluate the concordance (ie, agreement) between Web-based self-reported anthropometrics and face-to-face declaration. We hypothesized that (1) we would observe underreporting of BMI with the Web-based questionnaire compared with the gold standard (ie, clinical measurement), and (2) social desirability in front of the computer would be less important than on the phone compared with the face-to-face interview.

Methods

The NutriNet-Santé Study

The present analyses were carried out on a subsample of the NutriNet-Santé study, an ongoing Web-based prospective cohort study launched in France in May 2009 [28] aiming to investigate the associations between nutrition and health and to study the role of various determinants (sociodemographic, economic, biochemical, cognitive, etc) of dietary behavior and nutritional status. Recruitment of adult volunteers (aged ≥ 18 years) through multimedia campaigns is to be carried out for 5 years with a planned additional follow-up of 10 years.

Briefly, at inception, participants complete a set of Web-based questionnaires assessing socioeconomic and sociodemographic conditions, dietary intake, physical activity, anthropometrics, lifestyle, and health status [28]. Each month, participants are invited to fill in complementary optional questionnaires related to determinants of dietary behavior and nutritional and health status. The anthropometric questionnaire is repeated every 6 months.

Moreover, participants are invited to attend one of the specific health centers involved in the study, located in various French cities. During the visit, they undergo blood and urine sampling and a clinical examination including anthropometric measurements. Height is measured by a trained technician with a wall-mounted stadiometer without shoes to the nearest 0.5 cm [29]. Weight is measured with a calibrated scale (body composition analyzer BC-418MA, TANITA, Tokyo, Japan) to the nearest 0.1 kg, with participants wearing indoor clothes, without shoes, socks, or stockings. Height is entered manually into the TANITA software, and then weight is measured, with the data sent automatically to the database through a secured interface. Results are checked with the participant allowing for detection of any typing errors regarding height. Complete

information about the NutriNet-Santé study design can be found elsewhere [28].

This study was approved by the International Research Board of the French Institute for Health and Medical Research (IRB Inserm no: 0000388FWA00005831) and the French National Information and Citizen Freedom Committee (CNIL no: 908450 and no: 909216). The collection of biological samples and clinical data was approved by the Consultation Committee for the Protection of Participants in Biomedical Research (C09-42 on May 5, 2010) and the French National Information and Citizen Freedom Committee (CNIL no: 1460707).

Validation and Concordance Samples

To validate the self-reported anthropometrics, a random subsample of the participants with a scheduled clinical examination were invited to fill in a Web-based anthropometric questionnaire 3 days before their appointment at the health center. This minimizes weight variations because of a long time lag between reported and measured weight. The validation study started in November 2011 and ended in July 2012. All participants with a scheduled visit in this time range were invited to fill in the anthropometric questionnaire. A total of 2513 participants completed the questionnaire 3 days before and had attended the subsequent clinical visit. This constitutes the validation sample.

Among them, some randomly assigned participants were asked by the trained technicians to declare their height and weight on the day of the examination, before being measured. The concordance study started in February 2012. By July 2012, a total of 815 participants had provided Web-based weight and height 3 days before and in a face-to-face interview, constituting the concordance sample. We chose to stop inclusions and start the analyses in July 2012 because it provided a good balance between an acceptable sample size as reviewed [5] and a reasonable study duration.

Covariates

Socioeconomic variables were collected at study baseline. Education referred to the highest achieved level (primary school, secondary school, high school diploma, university bachelor degree or less, university graduates with higher than bachelor degree) and was further regrouped into 3 categories (up to high school diploma, university bachelor degree or less, university graduates with higher than bachelor degree); occupational category was defined according to the current job or the last job held for unemployed or retired individuals (never employed, self-employed, farmers, manual workers, intermediate professions, managerial/professional staff). Monthly household income and household composition (marital status, number and age of children) were also reported, which allowed calculating monthly income per household unit (in euros) by using a standardized algorithm [30] and were categorized in quartiles. Tobacco use (current, former, never smoked), and marital status were also used as covariates.

Leisure time physical activity (LTPA) was assessed by the International Physical Activity Questionnaire (IPAQ) [31,32] and classes of physical activity were defined as recommended [33] in low, medium, and high LTPA categories. LTPA data

are collected each year in the NutriNet-Santé study, so the most recent report was used.

Statistical Analysis

For comparison to self-declared data, measured weight was rounded to the nearest kilogram and height to the nearest centimeter. Log-transformation was applied to height, weight, and resulting BMI to improve normality. BMI was categorized as normal ($BMI < 25 \text{ kg/m}^2$), overweight excluding obesity ($BMI 25\text{--}29 \text{ kg/m}^2$), and obese ($BMI \geq 30 \text{ kg/m}^2$). Throughout this paper, *overweight* refers to overweight excluding obesity, unless otherwise stated.

Population characteristics (sex, age, socioeconomic status, tobacco use, LTPA, and anthropometrics) were compared between the validation and concordance samples and with the entire NutriNet-Santé cohort by *t* tests and chi-square (χ^2) tests.

A summary of the indicators used for validation and concordance analyses is provided in [Multimedia Appendix 1](#).

Validation Analysis

Several statistical procedures were used to assess the validity of Web-based self-reported anthropometrics by comparing them to the reference values measured by the technician. The difference between self-reported and measured weight, height, and resulting BMI were calculated. *P* value referred to paired *t* test (on log-transformed variables). To assess agreement between self-reported and measured values, a random effect model was performed to estimate intraclass correlation coefficient (ICC) (2,1) as proposed by Shrout and Fleiss [34] using the SAS macro %INTRACC [35]. We also used the Bland-Altman method [36]: for each variable (log transformed), the difference self-reported minus measured was plotted against the average (self-reported + measured)/2, providing mean agreement and 95% limits of agreement (LOA) defined as mean agreement ± 2 SD of the difference. Because results were antilogged after analysis, the mean agreement and LOA are given as ratio of self-reported to measured values [37,38]. A mean agreement of 100% represents exact agreement, otherwise there is systematic bias. If agreement is $>100\%$, it indicates that, on average, participants overreported, whereas $<100\%$ indicates underreporting compared to the measure. The slope of average of methods regressed on the difference between methods was also estimated to test the existence of proportional bias although the Bland-Altman method does not adequately distinguish between fixed and proportional bias [39]. To further investigate the influence of socioeconomic and lifestyle factors (BMI category, age, sex, LTPA, smoking status, education level, level of income, occupation), bivariate and multivariate regression analyses were used, considering the difference between self-reported and measured height, weight, or BMI as the dependent variables.

Percentage of agreement between self-reported and measured categories of BMI were calculated and the degree of misclassification was assessed through weighted kappa coefficient. McNemar tests were carried out for the binary variables (1) overweight including obesity and (2) obese. Sensitivity and specificity for overweight and obese were also

calculated as true positives/(true positives + false negatives) and true negatives/(true negatives + false positives), with the true measure being the clinical data.

Concordance Analyses

The same procedures were used for the concordance study between self-reported Web-based questionnaire and face-to-face interview, namely paired *t* test of the difference between Web and face-to-face values, ICC, Bland–Altman regression and LOA, percentage of agreement, and weighted kappa coefficient.

Sensitivity Analyses

Because participants who answered the Web-based anthropometric questionnaire 3 days before attending the visit knew that they would be measured, this could lead to overagreement between self-reported and measured data. To overcome this potential bias, we performed the following sensitivity analyses: a second validity sample included participants who filled in the regular Web-based anthropometric questionnaire (available every 6 months) within 2 months before attending the visit. The visit was not necessarily scheduled at time of completion; hence, participants were unaware of an upcoming measurement. A time lag of a maximum 2 months was chosen to limit actual weight variations. The second validity sample consisted of 2078 participants. Among them, a second concordance sample of 233 participants was drawn that had available data from the face-to-face declaration.

All statistical tests were 2-sided and $P < .05$ was considered significant. All statistical analyses were performed using SAS software ver 9.1 (SAS Institute Inc, Cary, NC, USA).

Results

Population Characteristics

The characteristics of the entire NutriNet-Santé cohort and of the validity and concordance samples are presented in [Table 1](#). There were no significant differences between the validity and concordance samples regarding age, education, occupation, smoking status, and LTPA. Participants in the validity sample were less often women, significantly older, more physically active, less likely to be smokers, and more likely to live with a partner and to have a higher level of income than the entire cohort (all P values $< .001$). Web-based self-reported anthropometrics showed no significant difference between the validity sample and the cohort, except for a slightly higher height ($P = .003$).

Validity

Men and women underreported their weight by -0.40 kg (SD 1.45) and -0.52 kg (SD 1.42), respectively, and overreported their height by 0.61 cm (SD 1.40) and 0.55 cm (SD 2.66), leading to an underreporting of BMI of -0.32 kg/m² (SD 0.66) for men and -0.34 kg/m² (SD 1.67) for women (all $P < .001$) ([Multimedia Appendix 2](#)). No difference was observed between

men and women for BMI, height (*t* test P values $> .05$), and weight ($P = .05$).

Validity of continuous variables is presented in [Table 2](#). Overall, agreement was high between self-reported and measured anthropometric data with ICC ranging from 0.94 (height) to 0.99 (weight). However, a systematic bias was observed for each variable because percent mean agreement was significantly different from 100%, indicating underreporting of weight and BMI and overreporting of height. The LOA were wider for BMI than for height and weight. For approximately 95% of cases, self-reported BMI differed from measured BMI by 8.9% less than to 6.7% greater than the real value (LOA are provided compared to the reference, ie, 100%, but are symmetrical in relation to the mean of agreement value, here 98.6%; [Figure 1](#)).

To investigate determinants of differential bias, we regressed the difference between self-reported and measured BMI values on covariates. BMI category showed a significant effect (crude and adjusted for covariates: sex, age, LTPA, occupation, education, and smoking). BMI underreporting was -0.16 , -0.36 , and -0.63 kg/m² among normal, overweight, and obese participants, respectively, in the adjusted model. Weight underreporting was significantly associated with BMI category (more underreporting among obese and overweight vs normal) and sex (women underreported more than men). Height overreporting was positively associated with BMI category (more overreporting among obese and overweight vs normal) and age. Crude differences by sex, across BMI and age categories are reported in [Multimedia Appendix 2](#).

[Table 3](#) shows an agreement of 93.2% between BMI categories and a weighted kappa of 0.89. The overweight proportion was 2.2 percentage points less when estimated from self-reported than from clinical data (23.5% vs 25.7%) and 1.7 points less for obesity (9.1% vs 10.7%). The difference was statistically significant according to the McNemar test ($P < .001$). Regarding detection of obesity, out of 270 truly obese participants, 45 were not classified obese with the self-report (false negative) whereas 225 were well-detected (true positive), leading to a sensitivity of 83.3% and a specificity of 99.9%. Regarding detection of overweight including obesity (BMI ≥ 25), 97 participants were false negative and 818 true positive, leading to a sensitivity of 87.9% and a specificity of 99.1%.

Concordance

As shown in [Table 4](#), mean agreement between Web-based and face-to-face values was almost perfect; the difference was not significant and ICCs were 1.00, 0.96, and 0.98 for weight, height, and BMI, respectively.

As presented in [Table 5](#), agreement in BMI categories was also very strong with 98.5% of the participants similarly classified in BMI classes. The weighted kappa was 0.97 and difference in overweight classification was not significant, but it was significant for obesity ($P = .01$).

Table 1. Characteristics of the validation study sample (N=2513) and the concordance study sample (n=815) from the NutriNet-Santé Study, 2012, France.

Participants' characteristics	Validity sample (V) n=2513	Concordance sample (C) ^a	NutriNet-Santé cohort (CO) n=115,784	P value ^b	
				V vs CO	C vs CO
Age (years), mean (SD)	53.8 (13.3)	53.6 (13.0)	45.1 (14.5)	<.001	<.001
Weight (kg)^c					
Mean (SD)	66.8 (13.2)	66.5 (13.4)	67.3 (15.1)	.06	.11
Median (IQR)	65 (57-75)	64 (57-74)	64 (57-75)		
Height (cm)^c					
Mean (SD)	166.3 (8.3)	165.7 (8.5)	166.8 (8.5)	.003	.001
Median (IQR)	165 (160-172)	165 (160-170)	166 (161-172)		
BMI (kg/m²)^c					
Mean (SD)	24.1 (4.3)	24.2 (4.4)	24.2 (5.2)	.49	.95
Median (IQR)	23.3 (21.1-26)	23.5 (21.2-26.2)	23.1 (20.8-26.2)		
Female, n (%)	1835 (73.0)	606 (74.4)	90,382 (78.1)	<.001	.01
Living with a partner, n (%)	1860 (74.0)	607 (74.5)	82,480 (71.2)	.001	.04
BMI (kg/m²),^d n (%)					
Normal (<25 kg/m ²)	1604 (63.8)	513 (62.9)	76,879 (67.2)		
Overweight (25-29 kg/m ²)	643 (25.6)	210 (25.8)	25,396 (22.2)		
Obese (≥30 kg/m ²)	266 (10.6)	92 (11.3)	12,125 (10.6)		
Education, n (%)					
Primary school	78 (3.2)	22 (2.8)	3854 (3.4)	.44	.96
Secondary school	491 (20.1)	156 (19.8)	19,971 (17.6)		
High school diploma	374 (15.3)	113 (14.3)	20,557 (18.1)		
University < bachelor degree	746 (30.5)	264 (33.4)	33,362 (29.5)		
University ≥ bachelor degree	757 (31.0)	235 (29.8)	35,552 (31.4)		
Occupational category, n (%)					
Never employed	55 (2.2)	18 (2.2)	6646 (5.7)	.39	.94
Self-employed, farmers	101 (4.0)	33 (4.1)	3951 (3.4)		
Manual workers	53 (2.1)	21 (2.6)	3509 (3.0)		
Intermediate professions	1372 (54.6)	436 (53.5)	65,223 (56.3)		
Managerial/professional	932 (37.1)	307 (37.7)	36,455 (31.5)		
Tobacco smoking, n (%)					
Current smoker	241 (9.6)	86 (10.5)	2079 (18.0)	<.001	.001
Former smoker	999 (39.7)	320 (39.3)	38,324 (33.1)		
Never smoker	1273 (50.7)	409 (50.2)	5667 (48.9)		
Physical activity level,^d n (%)					
Low	498 (20.3)	176 (22.1)	27,212 (25.6)	<.001	.79
Medium	1002 (40.8)	300 (37.6)	44,239 (41.7)		
High	954 (38.9)	322 (40.3)	34,695 (32.7)		
Level of income (€unit of consumption), n (%)					
				<.001	<.001

Participants' characteristics	Validity sample (V) n=2513	Concordance sample (C) ^a	NutriNet-Santé cohort (CO) n=115,784	P value ^b	
				V vs CO	C vs CO
Don't want to answer	261 (10.4)	91 (11.2)	14,929 (13.5)		
<1257	302 (12.0)	112 (13.7)	23,511 (21.3)		
1257-1835	508 (20.2)	166 (20.4)	23,606 (21.4)		
1835-2700	674 (26.8)	225 (27.6)	24,329 (22.1)		
>2700	768 (30.6)	221 (27.1)	23,849 (21.6)		

^aNo significant difference was observed between the validity and concordance samples (all *P* values >.05 for chi-square tests or *t* test for age).

^b*P* value for *t* test or Mantel–Haenszel chi-square test as appropriate.

^c*t* tests on the log-transformed variables.

^dReduced sample size because of missing values; validity sample: n=2454 for physical activity level; concordance sample: n=798 for physical activity level; cohort: n=114,400 for BMI, n=113,296 for education, n=106,146 for physical activity level.

Table 2. Validity indicators of weight, height, and body mass index (BMI) including intraclass correlation coefficient (ICC) between the Web-based self-report and measurement at the clinical examination, Bland–Altman mean agreement, and limits of agreement (LOA) from the NutriNet-Santé Study, 2012, France (N=2513).

Anthropometric variables	Web-based		Measured		Difference		<i>P</i> ^a	ICC ^b		% mean agreement ^c		% LOA ^d	
	Mean	SD	Mean	SD	Mean	SD		ICC	95% CI	%	95% CI	Lower limit	Upper limit
Weight (kg)	66.84	13.60	67.33	13.74	-0.49	1.43	<.001	0.99	0.99, 0.99	99.28	99.20, 99.37	95.11	103.64
Height (cm)	166.30	8.48	165.73	8.32	0.56	2.39	<.001	0.94	0.94, 0.95	100.33	100.27, 100.40	97.06	103.72
BMI (kg/m ²)	24.12	4.44	24.46	4.41	-0.34	1.47	<.001	0.97	0.97, 0.97	98.61	98.47, 98.77	91.12	106.74

^a*P* value of the paired *t* test of difference of log-transformed variable.

^bICC(2,1) calculated on log-transformed variables.

^cBland–Altman mean agreement (average of difference self-reported – measured). A mean agreement of 100% represents exact agreement between the 2 methods.

^dLOA: limits of agreement of self-reported value expressed as a percent of the measured value. Because results were antilogged after analysis, the LOA are given as ratio Web:measured.

Table 3. Validity indicators for categorical variables including percent of similar classification and weighted kappa coefficient for overweight and obesity classification between the Web-based declaration and reference measurement at clinical examination from the NutriNet-Santé Study, 2012, France (N=2513).

Categorical anthropometric variable	Web-based n=2513		Measured n=2513		Agreement (%)		Weighted kappa ^a		<i>P</i> ^b	Sensitivity ^{c,d}		Specificity ^{c,e}	
	n	%	n	%	%	95% CI	κ	95% CI		%	95% CI	%	95% CI
BMI classification					93.2	92.2, 94.1	0.89	0.88, 0.91					
Normal (BMI<25)	1695	67.45	1598	63.59									
Overweight (BMI 25-29.9)	590	23.48	645	25.67					<.001	87.9	0.86, 0.90	99.1	98.7, 99.6
Obese (BMI≥30)	228	9.07	270	10.74					<.001	83.3	78.9, 87.8	99.9	99.7, 100

^aCicchetti–Allison weight. For a given cell in row *i*, column *j*, $w_{ij}=1-(|i-j|/2)$.

^b*P* value of McNemar chi-square test for binary variables: overweight including obesity (BMI≥25) yes/no and obese (BMI≥30) yes/no. A *P* value <.05 indicates significant difference between Web-based self-reporting and measurement.

^cSensitivity and specificity for binary variables: overweight including obesity (BMI≥25) and obese (BMI≥30).

^dSensitivity=true positives/(true positives + false negatives).

^eSpecificity=true negatives/(true negatives + false positives). True = clinical data.

Figure 1. Bland - Altman plot of self-reported versus measured values of BMI, NutriNet-Santé study, 2012, France. Horizontal lines represent the % mean difference and 95% limits of agreement.

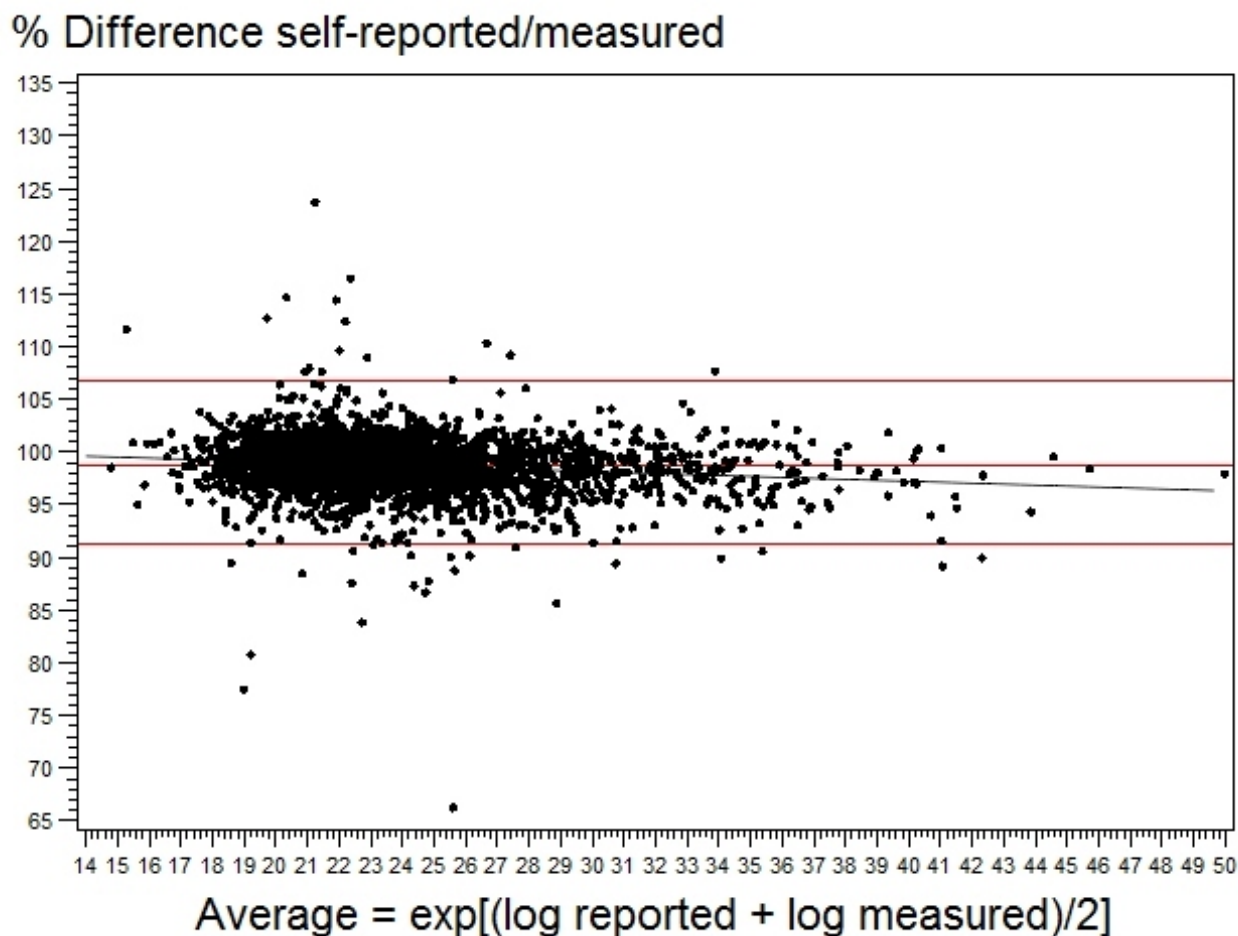


Table 4. Concordance indicators for continuous variables including intraclass correlation coefficient (ICC) between Web-based and face-to-face reported data, Bland–Altman mean agreement, and limits of agreement (LOA) from the NutriNet-Santé Study, 2012, France (n=815).

Anthropometric variable	Web-based		Face-to-face		Difference		<i>P</i> ^a	ICC ^b		% mean agreement ^c		% LOA ^d	
	Mean	SD	Mean	SD	Mean	SD		ICC	95% CI	%	95% CI	Lower limit	Upper limit
Weight (kg)	66.60	13.45	66.60	13.49	0.00	1.14	.31	0.996	0.995, 0.996	100.01	99.89, 100.14	96.46	103.69
Height (cm)	165.75	8.50	165.71	8.24	0.04	2.21	.77	0.958	0.951, 0.963	100.02	99.91, 100.12	97.13	102.98
BMI (kg/m ²)	24.20	4.40	24.19	4.28	0.01	1.20	.78	0.979	0.976, 0.982	100.00	99.80, 100.27	93.43	107.10

^a*P* value of the paired *t* test of difference of log-transformed variable (Web minus face-to-face).

^bICC: intraclass correlation (2,1) calculated on log-transformed variables.

^cBland and Altman mean agreement (average of differences “Web-based minus face-to-face”). A mean agreement of 100% represents exact agreement between the 2 questionnaires.

^dLOA: limits of agreement of Web-based self-reported value expressed as a percent of the face-to-face reported value. Because results were antilogged after analysis, the LOA are given as ratio Web-based/face-to-face.

Table 5. Concordance indicators for categorical variables: percent of similar classification and weighted Kappa coefficient for overweight and obesity classification between Web-based and face-to-face reported data from the NutriNet-Santé Study, 2012, France (n=815).

Categorical anthropometric variable	Web-based		Face-to-face		Agreement (%)		Weighted kappa ^a		P ^b
	n	%	n	%	%	95% CI	κ	95% CI	
BMI classification					98.5	97.7, 99.4	0.97	0.96, 0.99	
Normal (BMI<25)	547	67.1	546	67.0					
Overweight (BMI 25-29.9)	193	23.7	188	23.1					1.00
Obese (BMI≥30)	75	9.2	81	9.9					.01

^aCicchetti–Allison weight. For a given cell in row *i*, column *j*, $w_{ij}=1-(|i-j|/2)$

^bP value of McNemar chi-square test for binary variables: overweight including obesity (BMI>=25) yes/no and obese (BMI>=30) yes/no.

Sensitivity Analyses

Sensitivity analyses in the second validity sample (n=2078) showed similar results as the validity sample, the validity indicators (ICC, kappa, percent agreement) were even slightly higher (Multimedia Appendix 3). However, a significant difference in weight reporting was observed in the second concordance sample (n=233): participants reported higher weight (mean 0.37, SD 1.86) and, hence, BMI (mean 0.32, SD 0.83) in the Web-based questionnaire than in the face-to-face interview. Weighted kappa was lower than the concordance sample, with a value of 0.91 (95% CI 0.86-0.95) and percent correct classification was 94%. Nevertheless, ICCs were similar, ranging from 0.98 to 0.99.

Discussion

Principal Finding

In the present study, we observed that Web-based self-report of anthropometrics in the NutriNet-Santé study is equivalent to a face-to-face interview. Although, as hypothesized, it is subject to bias as compared with direct measures, the bias is reasonably small and the validity indicators show good reliability of this data.

Validity

Overall, our results showed high validity of self-reported anthropometric data compared with measured values. However, we observed a small although significant underreporting of weight and BMI and an overreporting of height, which was expected and is consistent with previous research [5]. Compared with the bias reported in the literature, the extent of misreporting in the present study (-0.49 kg for weight and 0.56 cm for height) is smaller than in most of the studies on general adult populations which show underreporting ranging from -0.1 to -6.5 kg for weight and overreporting from 0.6 to 7.5 cm for height [5]. Results of the other study assessing validity of Web-based self-reported anthropometrics showed greater underreporting of weight (-1.2 kg) and found no significant difference between men and women [27]. BMI classification is more of a concern when studying the association of nutritional factors with obesity or overweight risk. But, a correct classification of 93% and a kappa of 0.89 (which can be considered almost perfect [40]) reflect reliable and suitable results. For example, in the Adventist Health Study, correct

classification in BMI categories was 83.4% (95% CI 80.9%, 85.8%) [41]. In our study, sensitivity of self-reported BMI to detect obesity was 83% and specificity was 100%, which are higher than the Adventist study (sensitivity 81%, specificity 97%), and much higher than observed in a Swiss and French community-based sample (sensitivity: 66% for men, 73% for women; specificity: 99% for both) [16]. Regarding comparability of our study population with other studies, in the Adventist study, the prevalence of self-reported obesity (27.3%) was higher than in our study (9.1%); however, the study by Dauphinot et al [16] reported exactly the same proportion of obese participants as in our study.

No difference in misreporting was observed between men and women for height, whereas it has been previously suggested that men tended to overreport their height more strongly than women [8,10,14-16,42,43], although a few studies found no difference according to sex [9,27,41]. However, we found that being a woman was a predictor of greater underreporting of weight, consistent with previous research [8,10,14-16,38,39]. Age was a significant predictor for overreporting of height, in accordance with most the studies [9,11,41,44,45]. This can be, at least partly, explained by the fact that aging is associated with a decrease in height that people might be unaware of if they are not often measured [13,19].

Although underreporting of BMI and weight and overreporting of height was observed in every BMI category, their magnitude differed and we found that objective overweight and obesity were the strongest predictors for underreporting of weight and BMI and overreporting of height, similar to many studies [10,11,16,19,41,44,45]. Our results are very similar to those of the Adventist study [41] that showed a BMI underreporting of -0.4 kg/m² in nonobese vs -0.9 kg/m² in obese participants. We found lower differences between BMI categories than in the Oxford EPIC study [10] in which underreporting among normal, overweight, and obese participants was -0.6, -1.02, and -1.66 kg/m² for men and -0.44, -0.96, and -1.35 kg/m² for women, and in the study by McAdams et al [45] in which BMI misreporting was 0.03 (nonsignificant), -0.57, and -1.77 kg/m² in normal, overweight, and obese participants, respectively. Regarding weight underreporting, our results show less difference between BMI categories than Bonn et al [27] who found underreporting of -0.9 kg in participants with BMI<25 vs -2.1 kg in overweight/obese participants. A hypothesis to

explain this phenomenon lies in the social desirability concept: people are influenced by their desire to conform to perceived societal norms, and this is more important in obese participants [19].

Concordance

Method of data collection can influence responses to surveys [46]. Several studies reported stronger underestimation of weight and BMI with telephone reporting than with face-to-face interviews [46-48]. Some hypotheses have been proposed to interpret such findings [18], including the idea that social desirability may influence reporting that cannot be visually verified [48].

Contrarily, and as hypothesized, in our study we showed almost perfect agreement between the Web-based reporting and the face-to-face interview, arguing that behind the computer screen, participants do not seem more prone to social desirability bias. This can be explained by the greater feeling of anonymity on the Web than on the telephone [48], in which the involvement is greater when the interviewer is a person rather than a computer screen. Indeed, even if the participants knew they would be weighed and measured after the face-to-face interview, this did not appear to influence what they declared.

We were aware that the Web-based reporting might be partly biased because participants theoretically knew they would be weighed a few days later; thus, limiting prevarication bias. However, the sensitivity analysis provided similar results, with even higher values of Web-based weight vs face-to-face, closer to the true measure. This shows an advantage of Web-based self-report compared with telephone interview as we previously demonstrated concerning dietary data [49].

Strength and Limitations

The first limitation pertains to a potential underestimation of the difference between Web-based reports and measures because participants in our study knew they would attend the visit 3 days after filling in the Web-based questionnaire. However, the sensitivity analyses with data collected within 2 months before the visit showed similar results—even slightly higher validity—indicating that the difference seems not to be reduced by awareness of the upcoming examination.

Second, caution is also advised regarding the generalizability of our results. Indeed, the participants of the NutriNet-Santé study were recruited on a voluntary basis, implying that they might be particularly likely to engage in healthy behaviors; thus, a self-selection bias could have occurred in our population as in most prospective cohort studies. In particular, participants

were invited to answer an anthropometric questionnaire twice a year, so they were likely to be more aware of their true weight. Further, the present validation study is subject to an additional selection bias related to the participation to the visit because some characteristics, such as age, smoking status, or LTPA, were significantly different between the validation sample and the entire cohort. However, even if some socioeconomic characteristics were different, educational level, occupation, and the main outcomes, anthropometric values, were not significantly different of the entire cohort. Also, among the participants who attended the clinical examination, those participating in the face-to-face interview were randomly allocated.

A major strength of this validation study is its originality. This is the second study assessing validity of anthropometric data collected through a Web-based tool, but we used a wider range of statistical tools that allowed analyzing the validity in more depth on a wider sample than in the recently published study [27]. This type of study is of major interest with the arising development of e-epidemiology. Also, the sample size is large and ranks among the larger validation samples published [5]. Another great strength is that the elapsed time between Web-based self-report and direct measure was controlled for, equal for every participant, and sufficiently short to avoid any true potential change in weight. Moreover, the gold standard used here, measured weight and height, was obtained through a standardized protocol by a trained technician and data were sent directly through a secured interface to the database, avoiding any data entry mistakes. Finally, statistical analysis was not limited to correlation coefficients calculation, but acknowledged statistical tools for validation and concordance analysis were used [34,36,37,50,51].

In conclusion, this study indicates that Web-based weight and height data from the NutriNet-Santé study can be considered as valid enough to be used when studying associations of nutritional factors with anthropometric and health outcomes. However, underreporting of weight and BMI and overreporting of height was stronger among overweight and obese and we showed misclassification of overweight (sensitivity 87.8%) and obesity (sensitivity 83.3%) which leads us to advise caution when overweight and obesity are the main outcomes. Although it is subject to biases inherent to self-reported anthropometric measurements, the magnitude of such biases can be considered comparable to face-to-face interviews. Therefore, Web-based self-reported data appear to be an accurate and useful tool to assess anthropometric data.

Acknowledgments

This work was supported by grants from the Région Ile de France (CORDDIM) and Fondation Coeur et Artères.

The NutriNet-Santé study is supported and has received grants by the following institutions: Ministère de la Santé (DGS), Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Éducation pour la Santé (INPES), Fondation pour la Recherche Médicale (FRM), Institut de Recherche en Santé Publique (IRESP), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM), and Université Paris 13.

We thank all the scientists who carried out the NutriNet-Santé study and their teams, technicians, and assistants, who assisted their work. We especially thank Laurent Bourhis, data manager; Yasmina Chelghoum, Paul Flanzky, and Mohand Ait Oufella, computer scientists; and Karine Prevost and Nabil Cherifi, technicians.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Statistical analyses for validity and concordance of anthropometrics, NutriNet Santé Study, France, 2012.

[[PDF File \(Adobe PDF File\), 7KB - jmir_v15i8e152_app1.pdf](#)]

Multimedia Appendix 2

Difference between Web-based self-report and measured anthropometrics according to BMI classification and age class, by sex, NutriNet-Santé study, France 2012.

[[PDF File \(Adobe PDF File\), 37KB - jmir_v15i8e152_app2.pdf](#)]

Multimedia Appendix 3

Sensitivity analyses among subsample with a time lag between Web-based self-report and measurement < 2 months, NutriNet-Santé study, France, 2012.

[[PDF File \(Adobe PDF File\), 41KB - jmir_v15i8e152_app3.pdf](#)]

References

1. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser 2000;894:i-xii, 1. [Medline: [11234459](#)]
2. Joint WHO/FAO Expert Consultation. Technical Report Series No. 916. Geneva, Switzerland: WHO; 2003. Diet, Nutrition and the Prevention of Chronic Disease URL: <http://www.who.int/dietphysicalactivity/publications/trs916/download/en/index.html> [accessed 2013-07-16] [WebCite Cache ID [6IA16s3yc](#)]
3. World Health Organization Regional Office for Europe. The challenge of obesity in the WHO European Region and the strategies for response. Copenhagen, Denmark: WHO Regional Office for Europe; 2007. URL: http://www.euro.who.int/data/assets/pdf_file/0010/74746/E90711.pdf [accessed 2013-07-15] [WebCite Cache ID [6I8dZ2ENv](#)]
4. World Health Organization. Global Health Risks: Mortality and burden of disease attributable to selected major risks. Geneva, Switzerland: WHO; 2009. URL: http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf [accessed 2013-07-16] [WebCite Cache ID [6IA100iHc](#)]
5. Connor Gorber S, Tremblay M, Moher D, Gorber B. A comparison of direct vs self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* 2007 Jul;8(4):307-326. [doi: [10.1111/j.1467-789X.2007.00347.x](https://doi.org/10.1111/j.1467-789X.2007.00347.x)] [Medline: [17578381](#)]
6. Hill A, Roberts J. Body mass index: a comparison between self-reported and measured height and weight. *J Public Health Med* 1998 Jun;20(2):206-210 [FREE Full text] [Medline: [9675741](#)]
7. Kuskowska-Wolk A, Karlsson P, Stolt M, Rössner S. The predictive validity of body mass index based on self-reported weight and height. *Int J Obes* 1989;13(4):441-453. [Medline: [2793299](#)]
8. Niedhammer I, Bugel I, Bonenfant S, Goldberg M, Leclerc A. Validity of self-reported weight and height in the French GAZEL cohort. *Int J Obes Relat Metab Disord* 2000 Sep;24(9):1111-1118. [Medline: [11033979](#)]
9. Nyholm M, Gullberg B, Merlo J, Lundqvist-Persson C, Råstam L, Lindblad U. The validity of obesity based on self-reported weight and height: Implications for population studies. *Obesity (Silver Spring)* 2007 Jan;15(1):197-208. [doi: [10.1038/oby.2007.536](https://doi.org/10.1038/oby.2007.536)] [Medline: [17228048](#)]
10. Spencer EA, Appleby PN, Davey GK, Key TJ. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutr* 2002 Aug;5(4):561-565. [doi: [10.1079/PHN2001322](https://doi.org/10.1079/PHN2001322)] [Medline: [12186665](#)]
11. Villanueva EV. The validity of self-reported weight in US adults: a population based cross-sectional study. *BMC Public Health* 2001;1:11 [FREE Full text] [Medline: [11716792](#)]
12. Engstrom JL, Paterson SA, Doherty A, Trabulsi M, Speer KL. Accuracy of self-reported height and weight in women: an integrative review of the literature. *J Midwifery Womens Health* 2003;48(5):338-345. [Medline: [14526347](#)]
13. Rowland ML. Self-reported weight and height. *Am J Clin Nutr* 1990 Dec;52(6):1125-1133 [FREE Full text] [Medline: [2239790](#)]

14. Merrill RM, Richardson JS. Validity of self-reported height, weight, and body mass index: findings from the National Health and Nutrition Examination Survey, 2001-2006. *Prev Chronic Dis* 2009 Oct;6(4):A121 [FREE Full text] [Medline: 19754997]
15. Akhtar-Danesh N, Dehghan M, Merchant AT, Rainey JA. Validity of self-reported height and weight for measuring prevalence of obesity. *Open Med* 2008;2(3):e83-e88 [FREE Full text] [Medline: 21602953]
16. Dauphinot V, Wolff H, Naudin F, Guéguen R, Sermet C, Gaspoz JM, et al. New obesity body mass index threshold for self-reported data. *J Epidemiol Community Health* 2009 Feb;63(2):128-132. [doi: 10.1136/jech.2008.077800] [Medline: 18801799]
17. Kuskowska-Wolk A, Bergström R, Boström G. Relationship between questionnaire data and medical records of height, weight and body mass index. *Int J Obes Relat Metab Disord* 1992 Jan;16(1):1-9. [Medline: 1314237]
18. St-Pierre M, Béland Y. Mode effects in the Canadian Community Health Survey: a Comparison of CAPI and CATI. 2004. URL: http://www23.statcan.gc.ca/imdb-bmdi/document/3226_D16_T9_V5-eng.htm [accessed 2013-07-15] [WebCite Cache ID 6IA1tCGxn]
19. Larson MR. Social desirability and self-reported weight and height. *Int J Obes Relat Metab Disord* 2000 May;24(5):663-665. [Medline: 10849592]
20. Ekman A, Klint A, Dickman PW, Adami HO, Litton JE. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. *Eur J Epidemiol* 2007;22(5):293-300. [doi: 10.1007/s10654-006-9091-0] [Medline: 17206467]
21. Ekman A, Litton JE. New times, new needs; e-epidemiology. *Eur J Epidemiol* 2007;22(5):285-292. [doi: 10.1007/s10654-007-9119-0] [Medline: 17505896]
22. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. *Eur J Epidemiol* 2010 May;25(5):297-304 [FREE Full text] [doi: 10.1007/s10654-010-9431-y] [Medline: 20148289]
23. Schatzkin A, Subar AF, Moore S, Park Y, Potischman N, Thompson FE, et al. Observational epidemiologic studies of nutrition and cancer: the next generation (with better observation). *Cancer Epidemiol Biomarkers Prev* 2009 Apr;18(4):1026-1032 [FREE Full text] [doi: 10.1158/1055-9965.EPI-08-1129] [Medline: 19336550]
24. Smith B, Smith TC, Gray GC, Ryan MA, Millennium Cohort Study Team. When epidemiology meets the Internet: Web-based surveys in the Millennium Cohort Study. *Am J Epidemiol* 2007 Dec 1;166(11):1345-1354 [FREE Full text] [doi: 10.1093/aje/kwm212] [Medline: 17728269]
25. van Gelder MM, Breveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? *Am J Epidemiol* 2010 Dec 1;172(11):1292-1298 [FREE Full text] [doi: 10.1093/aje/kwq291] [Medline: 20880962]
26. Touvier M, Méjean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010 May;25(5):287-296. [doi: 10.1007/s10654-010-9433-9] [Medline: 20191377]
27. Bonn SE, Trolle Lagerros Y, Bälter K. How valid are Web-based self-reports of weight? *J Med Internet Res* 2013;15(4):e52 [FREE Full text] [doi: 10.2196/jmir.2393] [Medline: 23570956]
28. Herceberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 2010;10:242 [FREE Full text] [doi: 10.1186/1471-2458-10-242] [Medline: 20459807]
29. Lohman T, Roche A, Martorell R. Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics Books; 1988.
30. Institut national de la statistique et des études économiques (Insee) [National Institute of Statistics and Economic Studies]. Unités de consommation [consumption units] URL: <http://www.insee.fr/en/methodes/default.asp?page=definitions/unite-consommation.htm> [accessed 2013-07-15] [WebCite Cache ID 6I8f30XEe]
31. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr* 2006 Sep;9(6):755-762. [Medline: 16925881]
32. 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]
33. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005 Nov. URL: <http://www.ipaq.ki.se/scoring.pdf> [accessed 2013-07-15] [WebCite Cache ID 6I8f79XOL]
34. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979 Mar;86(2):420-428. [Medline: 18839484]
35. The SAS Institute. %INTRACC macro URL: http://support.sas.com/kb/25/addl/fusion25031_1_intracc.sas.txt [accessed 2013-07-15] [WebCite Cache ID 6I8fARaAO]
36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
37. Bland JM, Altman DG. Agreed statistics: measurement method comparison. *Anesthesiology* 2012 Jan;116(1):182-185 [FREE Full text] [doi: 10.1097/ALN.0b013e31823d7784] [Medline: 22129533]

38. Euser AM, Dekker FW, le Cessie S. A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. *J Clin Epidemiol* 2008 Oct;61(10):978-982. [doi: [10.1016/j.jclinepi.2007.11.003](https://doi.org/10.1016/j.jclinepi.2007.11.003)] [Medline: [18468854](https://pubmed.ncbi.nlm.nih.gov/18468854/)]
39. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. *Clin Exp Pharmacol Physiol* 2002 Jul;29(7):527-536. [Medline: [12060093](https://pubmed.ncbi.nlm.nih.gov/12060093/)]
40. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977 Jun;33(2):363-374. [Medline: [884196](https://pubmed.ncbi.nlm.nih.gov/884196/)]
41. Bes-Rastrollo M, Sabaté J, Jaceldo-Siegl K, Fraser GE. Validation of self-reported anthropometrics in the Adventist Health Study 2. *BMC Public Health* 2011;11:213 [FREE Full text] [doi: [10.1186/1471-2458-11-213](https://doi.org/10.1186/1471-2458-11-213)] [Medline: [21466678](https://pubmed.ncbi.nlm.nih.gov/21466678/)]
42. Bolton-Smith C, Woodward M, Tunstall-Pedoe H, Morrison C. Accuracy of the estimated prevalence of obesity from self reported height and weight in an adult Scottish population. *J Epidemiol Community Health* 2000 Feb;54(2):143-148 [FREE Full text] [Medline: [10715748](https://pubmed.ncbi.nlm.nih.gov/10715748/)]
43. Roberts RJ. Can self-reported data accurately describe the prevalence of overweight? *Public Health* 1995 Jul;109(4):275-284. [Medline: [7667492](https://pubmed.ncbi.nlm.nih.gov/7667492/)]
44. Kuczmarski MF, Kuczmarski RJ, Najjar M. Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. *J Am Diet Assoc* 2001 Jan;101(1):28-34; quiz 35. [doi: [10.1016/S0002-8223\(01\)00008-6](https://doi.org/10.1016/S0002-8223(01)00008-6)] [Medline: [11209581](https://pubmed.ncbi.nlm.nih.gov/11209581/)]
45. McAdams MA, Van Dam RM, Hu FB. Comparison of self-reported and measured BMI as correlates of disease markers in US adults. *Obesity (Silver Spring)* 2007 Jan;15(1):188-196. [doi: [10.1038/oby.2007.504](https://doi.org/10.1038/oby.2007.504)] [Medline: [17228047](https://pubmed.ncbi.nlm.nih.gov/17228047/)]
46. Hood K, Robling M, Ingledew D, Gillespie D, Greene G, Ivins R, et al. Mode of data elicitation, acquisition and response to surveys: a systematic review. *Health Technol Assess* 2012 May;16(27):1-162 [FREE Full text] [doi: [10.3310/hta16270](https://doi.org/10.3310/hta16270)] [Medline: [22640750](https://pubmed.ncbi.nlm.nih.gov/22640750/)]
47. Ezzati M, Martin H, Skjold S, Vander Hoorn S, Murray CJ. Trends in national and state-level obesity in the USA after correction for self-report bias: analysis of health surveys. *J R Soc Med* 2006 May;99(5):250-257 [FREE Full text] [doi: [10.1258/jrsm.99.5.250](https://doi.org/10.1258/jrsm.99.5.250)] [Medline: [16672759](https://pubmed.ncbi.nlm.nih.gov/16672759/)]
48. Joinson A. 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/)]
49. Touvier M, Kesse-Guyot E, Méjean C, Pollet C, Malon A, Castetbon K, et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011 Apr;105(7):1055-1064. [doi: [10.1017/S0007114510004617](https://doi.org/10.1017/S0007114510004617)] [Medline: [21080983](https://pubmed.ncbi.nlm.nih.gov/21080983/)]
50. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet* 1995 Oct 21;346(8982):1085-1087. [Medline: [7564793](https://pubmed.ncbi.nlm.nih.gov/7564793/)]
51. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011 Jan;64(1):96-106. [doi: [10.1016/j.jclinepi.2010.03.002](https://doi.org/10.1016/j.jclinepi.2010.03.002)] [Medline: [21130355](https://pubmed.ncbi.nlm.nih.gov/21130355/)]

Abbreviations

- BMI:** body mass index
- ICC:** intraclass correlation coefficient
- LOA:** limits of agreement
- LTPA:** leisure time physical activity

Edited by G Eysenbach; submitted 15.02.13; peer-reviewed by L Ware, J Nunez-Cordoba, B Brumen; comments to author 24.04.13; revised version received 23.05.13; accepted 31.05.13; published 08.08.13.

Please cite as:

Lassale C, Péneau S, Touvier M, Julia C, Galan P, Hercberg S, Kesse-Guyot E
Validity of Web-Based Self-Reported Weight and Height: Results of the Nutrinet-Santé Study
J Med Internet Res 2013;15(8):e152
URL: <http://www.jmir.org/2013/8/e152/>
doi: [10.2196/jmir.2575](https://doi.org/10.2196/jmir.2575)
PMID: [23928492](https://pubmed.ncbi.nlm.nih.gov/23928492/)

©Camille Lassale, Sandrine Péneau, Mathilde Touvier, Chantal Julia, Pilar Galan, Serge Hercberg, Emmanuelle Kesse-Guyot. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 08.08.2013. 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

Determinants of Participation in a Web-Based Health Risk Assessment and Consequences for Health Promotion Programs

Maurice AJ Niessen¹, MA; Eva L Laan², MSc; Suzan JW Robroek³, PhD; Marie-Louise Essink-Bot², PhD; Niels Peek⁴, PhD; Roderik A Kraaijenhagen¹, PhD; Coen K Van Kalken¹, PhD; Alex Burdorf³, PhD

¹NIPED Research Foundation, Amsterdam, Netherlands

²Department of Public Health, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

³Department of Public Health, Erasmus MC, Rotterdam, Netherlands

⁴Department of Medical Informatics, Academic Medical Center, Amsterdam, Netherlands

Corresponding Author:

Maurice AJ Niessen, MA

NIPED Research Foundation

Courbetstraat 34 H

Amsterdam, 1077 ZV

Netherlands

Phone: 31 20 2610444

Fax: 31 20 2610440

Email: m.a.j.niessen@niped.nl

Abstract

Background: The health risk assessment (HRA) is a type of health promotion program frequently offered at the workplace. Insight into the underlying determinants of participation is needed to evaluate and implement these interventions.

Objective: To analyze whether individual characteristics including demographics, health behavior, self-rated health, and work-related factors are associated with participation and nonparticipation in a Web-based HRA.

Methods: Determinants of participation and nonparticipation were investigated in a cross-sectional study among individuals employed at five Dutch organizations. Multivariate logistic regression was performed to identify determinants of participation and nonparticipation in the HRA after controlling for organization and all other variables.

Results: Of the 8431 employees who were invited, 31.9% (2686/8431) enrolled in the HRA. The online questionnaire was completed by 27.2% (1564/5745) of the nonparticipants. Determinants of participation were some periods of stress at home or work in the preceding year (OR 1.62, 95% CI 1.08-2.42), a decreasing number of weekdays on which at least 30 minutes were spent on moderate to vigorous physical activity (OR_{dayPA} 0.84, 95% CI 0.79-0.90), and increasing alcohol consumption. Determinants of nonparticipation were less-than-positive self-rated health (poor/very poor vs very good, OR 0.25, 95% CI 0.08-0.81) and tobacco use (at least weekly vs none, OR 0.65, 95% CI 0.46-0.90).

Conclusions: This study showed that with regard to isolated health behaviors (insufficient physical activity, excess alcohol consumption, and stress), those who could benefit most from the HRA were more likely to participate. However, tobacco users and those who rated their overall health as less than positive were less likely to participate. A strong communication strategy, with recruitment messages that take reasons for nonparticipation into account, could prove to be an essential tool for organizations trying to reach employees who are less likely to participate.

(*J Med Internet Res* 2013;15(8):e151) doi:[10.2196/jmir.2387](https://doi.org/10.2196/jmir.2387)

KEYWORDS

participation; Internet; workplace; health promotion; health risk assessment; reach

Introduction

Seven modifiable risk factors account for more than half of the chronic disease burden: high blood pressure, tobacco use, excess alcohol consumption, high serum cholesterol, overweight, low

fruit/vegetable intake, and physical inactivity [1]. The workplace is considered to be an excellent setting for health promotion programs that target these risk factors, not only because a large proportion of the population can be reached, but also because it makes use of a natural social network and can facilitate the

creation of a health-conscious environment [2-4]. Web-based interventions serve as a feasible and acceptable delivery method for these programs because they can provide scale at a relatively low cost per employee [5,6]. In addition, Internet access is available 24 hours a day, 7 days a week, which may serve both the employer and the employee, as program access is available across work shifts and into vacation and leisure time [6].

Recent reviews of effectiveness studies concluded there is sufficient evidence that worksite health promotion programs (WHPPs) have meaningful effects on a number of risk factors [7,8]. The latter is directly beneficial for the employer: implementing a WHPP can lead to reductions in both absenteeism and productivity loss at work [9,10]. However, a lack of employee participation presents an important barrier to the impact of WHPPs [7,11]. Since most intervention studies on WHPPs randomize workers who have agreed to participate in the studies, it is largely unknown whether those who could benefit most from the intervention are as likely to participate as those who may have already been making more healthful choices [12,13]. The importance of studying determinants of participation in WHPPs was already emphasized 25 years ago and has been underscored ever since [14-16]. Still, in 2009, the authors of a review concluded that few studies have evaluated the influence of health, lifestyle, and work-related factors on participation, which hampers insight into the underlying determinants of participation in WHPPs, and ultimately, the influence of selective participation on the effectiveness of these WHPPs [3]. Except for the finding that women enroll more often than men, no consistent determinants of participation in WHPPs aimed at physical activity and nutrition were found [3].

With regard to Web-based delivery of WHPPs, it has been reported that women and older people are more likely to enroll in these programs, as they more often use the Internet for searching for health-related information. It has also been postulated that individuals with a low educational level are less likely to use Web-based WHPPs, as those with less formal education are less likely to continue the adoption of innovations [17].

One type of WHPP that is frequently offered is the health risk assessment (HRA), which screens for risk factors for chronic diseases [7,10] and delivers verbal or written feedback on one's personal risk profile along with subsequent recommendations for lifestyle improvements. While an HRA is often used as a gateway intervention to broader WHPPs, it can also be utilized as a tool for stimulating the initiation of health behavior change [4,7]. In the current study, our aim was to analyze whether individual characteristics (including demographics, health behavior, self-rated health, and work-related factors) are associated with participation and nonparticipation in a Web-based HRA [9] implemented among employees in the Netherlands.

Methods

Participating Organizations and Study Design

In this cross-sectional study, the HRA was implemented in five Dutch organizations, which included a university medical center,

a large state-owned bank, a small bank, a financial institution, and the Dutch branch of an American multinational technology and consulting corporation. The HRA was applied in a pilot study among selected departments of the university medical center, which employed over 10,000 employees in 2009. The large state-owned bank was nationalized as a result of the global financial crises and employed more than 27,000 employees in 2009. Starting in 2006, its employees were gradually invited to enroll in the HRA. Renewed enrollment in the HRA was offered to employees 3 years after the first HRA was completed. In the current study, we included all invitees from 2009 who had not previously participated in the HRA. All workers from the small bank (<1000 employees) were invited, and from the financial institution (>3000 employees), all invitees from 2009 who had not previously participated in the HRA (renewed participation offered after 3 years) were included in this study. The Dutch branch of the American multinational technology and consulting corporation employed over 4500 employees in 2010. The HRA has been implemented in the organization since 2006. Two years after initial participation, renewed enrollment in the HRA is offered. In this study, we included all employees who were invited during the first and second quarters of 2010 and had not previously participated in the HRA.

Procedures

Employees were invited to participate in the HRA during the period from January 2009 to August 2010. The university medical center imposed an age criterion, inviting employees who were at least 45 years old. Upper management encouraged managers of selected departments to stimulate enrollment in the HRA among their workers. The HRA was also highlighted in the in-house employee magazine.

During the study period, invitations to participate in the HRA were sent by the human resources department, management, or the safety, health, and welfare services of the organizations involved. The invitation email included a description of the HRA and informed employees that participation was voluntary and free of charge, that all personal data would be treated confidentially, and that no individual results would be shared with their employer or any other party. No incentives were offered.

The HRA is called "The Prevention Compass" [4,9]. In the assessment phase, a Web-based health questionnaire is completed (in 30-45 minutes), biometric measurements (height, weight, waist circumference, blood pressure) are taken, and blood, urine, and feces samples are analyzed. A personalized Web-based health report and health plan is automatically generated only after all health data are collected. At this point, the HRA is completed.

Employees were defined as enrollees when they enrolled in the program by activating their online account during the inclusion period. This period varied (3-12 months), as larger organizations chose to invite their employees gradually. Enrollees who completed all HRA measurements within 1 year after the inclusion period had ended were classified as participants. Those who enrolled but did not complete all measurements were labeled dropouts. Employees who had not enrolled in the program after the inclusion period had ended were labeled

nonparticipants. The provider of the HRA sent nonparticipants an email inviting them to complete an online questionnaire. Those who responded to the online questionnaire were classified as responders, and those who did not respond were labeled nonresponders. Informed consent was obtained from all study participants prior to the study in accordance with the requirements for identifiable data collection in the Dutch Code of Conduct for Observational Research.

Measurements

For all study participants, gender and date of birth were available from the HRA invitation lists used by the organizations involved. Other individual characteristics (which included educational level, self-rated health, physical activity, body mass index (BMI), alcohol consumption, stress, work ability, and absenteeism during the previous year) were collected from the Web-based health questionnaire component of the HRA as part of a larger set of health data collected to generate a personal health report. As nonparticipants did not participate in the HRA and its Web-based health questionnaire, an online questionnaire was created that was made up almost entirely of the questions related to the above-mentioned individual characteristics of this study. Our goal was to lower the threshold and make it easier for nonparticipants to complete the questionnaire. Therefore, it was anonymous, no account had to be activated, and it took 10 minutes to complete. The questions relating to the individual characteristics were identical for participants and nonparticipants.

To determine educational level, respondents were asked to check 1 of 9 categories (ranging from no education to doctorate level) that indicated the highest level of education ever completed. Self-rated health [18,19] was measured by one question: "How do you rate your health in general?" The response options were "very good", "good", "moderate", "bad", or "very bad". Because of a lack of observations for the option "very bad", this category was merged with "bad" prior to the regression analysis.

One item derived from the Dutch version of the International Physical Activity Questionnaire [20] was used to assess the number of weekdays on which at least 30 minutes were spent on moderate to vigorous physical activity. BMI was based on height and weight as reported by respondents on the online questionnaire (nonparticipants) or measured by trained personnel (participants), and categorized into normal weight ($BMI < 25 \text{ kg/m}^2$), overweight ($25 \leq BMI < 30 \text{ kg/m}^2$), or obese ($BMI \geq 30 \text{ kg/m}^2$).

Alcohol consumption was measured in units of alcohol per week based on a standard alcohol questionnaire of the Dutch Municipal Health Service ("GGD Monitor"). Because few participants reported high levels of alcohol consumption, answer categories "29–35 units", "36–42 units", "43–50 units", and "> 50 units" were merged with "22–28 units" into "≥22 units." One item measured the frequency of tobacco use (none, occasionally, weekly, or daily). Answer categories "daily" and "weekly" were merged into "daily/weekly" as a measure of frequent tobacco use.

Items from the INTERHEART study were used to measure general and financial stress [21]. In accordance with the methods

used in that study, 2 items relating to stress at home and stress at work were combined into a general stress scale and graded as follows: (1) never experienced stress, (2) experienced some periods at home or at work, (3) experienced several periods at home or at work, or (4) experienced permanent stress at home or at work. Level of financial stress was defined as (1) little or none, (2) moderate, or (3) high or severe.

Work ability was measured with the single-item question on work ability from the Work Ability Index (WAI) [22]. Both the WAI and the single-item question show similar patterns of associations with absenteeism, health, and symptoms [23]. On the single-item question, respondents were asked to assess their current work ability compared with their lifetime best, with a possible score of 0 ("completely unable to work") to 10 ("work ability at its best").

Absenteeism during the previous 12-month period was determined by a question that classified the number of absenteeism (calendar) days related to health problems into 1 of 5 categories (0, 1-9, 10-24, 25-99, 100-365) [24].

Statistical Analysis

Means and standard deviations were presented for the continuous variables of age, physical activity, and work ability. Percentages were presented for the dichotomous variable gender and the categorical variables of education, BMI, alcohol consumption, tobacco use, stress at home or work, financial stress, self-rated health, and absenteeism. Enrollees, participants, nonparticipants, questionnaire responders, and nonresponders were compared using the unpaired *t* test for continuous variables and the chi-square test for dichotomous and categorical variables.

Spearman rank correlation coefficients were computed to investigate interrelationships among individual characteristics. Using the Bonferroni approach to control for Type 1 errors across the 132 correlations of the 12 variables, a *P* value of less than .0004 ($.05/132 = .0004$) was required for significance [25]. Correlations had to be at least 0.20 to be considered practically relevant.

Multivariate logistic regression analysis was performed to identify individual characteristics that contributed to participation in the HRA, after controlling for company and all other variables. This method presumes that all individual characteristics are measured for all cases and incomplete cases are discarded, which may result in biased estimates [26]. Therefore, multiple imputation of missing values of independent variables was employed. In multiple imputation, missing data are imputed based on variables correlated with the missing data and causes of missingness. In this study, ordinary least-squares regression models were applied to predict the missing values of continuous and ordinal variables, and discriminant prediction models were applied to the missing values of nominal variables. All individual characteristics as well as participant status (participant vs nonparticipant) were used as covariates in the predictive models. Uncertainty was accounted for by creating 10 imputed datasets [27]. Multivariate logistic regression analysis was carried out on each imputed dataset, producing multiple analysis results. These analysis results were combined

using rules established by Rubin [27] to produce one overall analysis, which is reported and compared with the results of complete case analysis.

The SOLAS 4 statistical package was used for the multiple imputation of the missing values. All other analyses were performed using SPSS for Windows, version 19.

Results

The study flow chart is presented in [Figure 1](#). During the study period, 8431 employees were invited to participate in the HRA. Average participation was 31.9% (2686/8431) and ranged from 14.9% to 51.7%: university medical center: 51.7% (206/503), state-owned bank: 29.9% (1282/4284), small bank: 41.0% (213/520), financial institution: 34.3% (824/2404), Dutch branch of American multinational technology and consulting corporation: 14.9% (107/720). The online questionnaire was completed by 27.2% (1564/5745) of the nonparticipants. Data on gender and age were available for 99.5% (8390/8431) of all HRA invitees from the invitation lists. Both enrollees ($P<.001$) and questionnaire responders ($P=.02$) were slightly older compared with nonparticipants and nonresponders. Also, enrollees were less often male ($P=.046$). Of those who enrolled in the HRA, 7.9% (213/2686) did not complete participation (dropouts). Compared with participants who completed the HRA, dropouts were younger ($P=.002$) and less often male ($P<.001$). Dropouts were excluded from further analysis, as no additional data beyond age and gender were available for this group. An example of a personal health risk profile page that was presented to those who completed the HRA is shown in [Figure 2](#).

[Table 1](#) depicts the baseline characteristics of participants (those who completed the HRA) and nonparticipants who filled in the online questionnaire (hereafter described as nonparticipants). Participants were slightly older than nonparticipants. No differences in gender or education were found. Participants engaged in physical activity less frequently, had higher weekly alcohol consumption, and reported having had periods of stress

at home or work during the previous year more often. Nonparticipants had lower self-rated health, used more tobacco, and reported slightly lower work ability, a higher level of financial stress, and more absenteeism in the preceding year.

A correlation matrix was computed to ascertain associations between the individual characteristics. Male gender was positively related with alcohol consumption ($r=.33$) and age was positively related with BMI ($r=.21$). A negative correlation ($r=-.28$) was found between the amount of stress at home or work and self-estimated work ability. Stress at home or work was positively correlated ($r=.21$) with financial stress. More positive self-rated health was correlated with higher work ability ($r=.29$) and negatively correlated with the amount of absenteeism during the previous 12-month period ($r=-.22$).

In [Table 2](#), the independent influence of demographics, health behavior, self-rated health, and work-related factors on HRA participation is shown for the imputed datasets (combined results), after controlling for organization (not shown) and all other independent variables. In the multivariate logistic regression analysis model, no effects were found for demographics. Less frequent physical activity, higher weekly alcohol consumption, and some periods of stress at home or work during the previous year remained statistically significantly associated with higher participation. It was also confirmed that less-than-positive self-rated health and tobacco use are significantly associated with lower participation. Higher levels of financial stress, more absenteeism, and lower work ability were no longer significantly related to lower participation.

Complete case analysis confirmed the direction of the reported results based on the imputed datasets. In addition, the following associations attained significance in the complete case analysis. Severe levels of financial stress, good self-rated health, and absenteeism (1-9 days and 100-365 days) were associated with lower participation. Having had several periods of stress at home or work and female gender were associated with higher participation. Also, in the complete case analysis, the association between occasional tobacco use and lower participation was marginally significant ($P=.06$).

Table 1. Baseline characteristics of HRA participants and nonparticipants who completed the online questionnaire.

Characteristics	HRA participants N=2473	HRA nonparticipants who completed questionnaire N=1564	P value
Age	n=2473	n=1564	.001
Mean (SD)	43.7 (9.2)	42.6 (9.7)	
Gender, n (%)	n=2472	n=1564	.81
Male	1337 (54.1)	852 (54.5)	
Female	1135 (45.9)	712 (45.5)	
Education^a, n (%)	n=2451	n=1549	.41
Low	400 (16.3)	266 (17.2)	
Intermediate	782 (31.9)	464 (30.0)	
High	1269 (51.8)	819 (52.9)	
Physical activity	n=2473	n=1403	<.001
Weekdays (0-7) ≥30 min., mean (SD)	3.2 (2.1)	3.8 (2.2)	
Body mass index (BMI), n (%)	n=2473	n=1404	.42
Normal weight: BMI <25kg/m ²	1078 (43.6)	586 (41.6)	
Overweight: BMI ≥25 - <30 kg/m ²	1097 (44.4)	637 (45.3)	
Obese: BMI ≥ 30 kg/m ²	298 (12.1)	184 (13.1)	
Alcohol consumption, n (%)	n=2473	n=1403	<.001
<1 units/wk	702 (28.4)	552 (39.3)	
1-7 units/wk	1037 (41.9)	569 (40.6)	
8-14 units/wk	479 (19.4)	195 (13.9)	
15-21 units/wk	173 (7.0)	64 (4.6)	
≥22 units/wk	82 (3.3)	23 (1.6)	
Tobacco use, n (%)	n=2471	n=1251	<.001
None	1961 (79.4)	889 (71.1)	
Occasional	115 (4.7)	79 (6.3)	
At least once/wk	395 (16.0)	283 (22.6)	
Stress—at home or work, n (%)	n=2436	n=1374	<.001
Never	278 (11.4)	194 (14.1)	
Some periods	1298 (53.3)	628 (45.7)	
Several periods	822 (33.7)	522 (38.0)	
Permanent	38 (1.6)	30 (2.2)	
Stress—financial, n (%)	n=2432	n=1374	<.001
Little or none	1872 (77.0)	947 (68.9)	
Moderate	490 (20.1)	352 (25.6)	
High or severe	70 (2.9)	75 (5.5)	
Self-rated health, n (%)	n=2468	n=1564	<.001
Very good	438 (17.7)	194 (12.4)	

Characteristics	HRA participants N=2473	HRA nonparticipants who completed questionnaire N=1564	<i>P</i> value
	Good	1684 (68.2)	1055 (67.5)
	Moderate	328 (13.3)	272 (17.4)
	Bad or very bad	18 (0.7)	43 (2.7)
Absenteeism, n (%)	n=2469	n=1374	<.001
	0 days	975 (39.5)	462 (33.6)
	1-9 days	1194 (48.4)	683 (49.7)
	10-24 days	183 (7.4)	117 (8.5)
	25-99 days	86 (3.5)	73 (5.3)
	100-365 days	31 (1.3)	39 (2.8)
Work ability	n=2466	n=1374	
	Mean (SD)	8.1 (1.4)	8.0 (1.5)
			.007

^aEducation: Low-lower general secondary/lower vocational; Intermediate-higher general secondary/pre-university/intermediate vocational; High-higher vocational/university.

Table 2. Influence of demographics, health, and work-related factors on HRA participation.

Characteristics		OR ^a	95% CI ^b
Age	10 yr intervals	1.127	0.961 - 1.322
Male gender		0.884	0.661 - 1.181
Education ^c	Low ^d		
	Intermediate	1.203	0.813 - 1.780
	High	0.919	0.618 - 1.365
Physical activity	Days per week ≥ 30 min.(0-7)	0.843	0.793 - 0.895
Body mass index (BMI)	Normal weight: BMI < 25 kg/m ^{2d}		
	Overweight: BMI ≥ 25 - < 30 kg/m ²	0.893	0.674 - 1.185
	Obese: BMI ≥ 30 kg/m ²	0.938	0.610 - 1.441
Alcohol consumption	< 1 units per week ^d		
	1-7 units per week	1.447	1.074 - 1.949
	8-14 units per week	1.971	1.318 - 2.947
	15-21 units per week	2.224	1.210 - 4.088
	≥ 22 units per week	3.372	1.317 - 8.632
Tobacco use	None ^d		
	Occasional	0.303	0.186 - 0.494
	At least once a week	0.645	0.461 - 0.903
Stress—home or work	Never ^d		
	Some periods	1.618	1.081 - 2.421
	Several periods	1.467	0.950 - 2.226
	Permanent	1.505	0.534 - 4.240
Stress—financial	Little or none ^d		
	Moderate	0.777	0.571 - 1.056
	High or severe	0.650	0.329 - 1.282
Self-rated health	Very good ^d		
	Good	0.711	0.489 - 1.035
	Moderate	0.567	0.344 - 0.935
	Bad or very bad	0.251	0.077 - 0.812
Absenteeism	0 days ^d		
	1-9 days	0.851	0.642 - 1.128
	10-24 days	0.719	0.442 - 1.172
	25-99 days	0.751	0.390 - 1.446
	100-365 days	0.480	0.177 - 1.302
Work ability	(0-10)	1.014	0.919 - 1.120

^aOR: odds ratio^bCI: confidence interval^cEducation: Low-lower general secondary/lower vocational; Intermediate-higher general secondary/pre-university/intermediate vocational; High-higher vocational/university.^dReference category

Figure 1. Study flow chart.

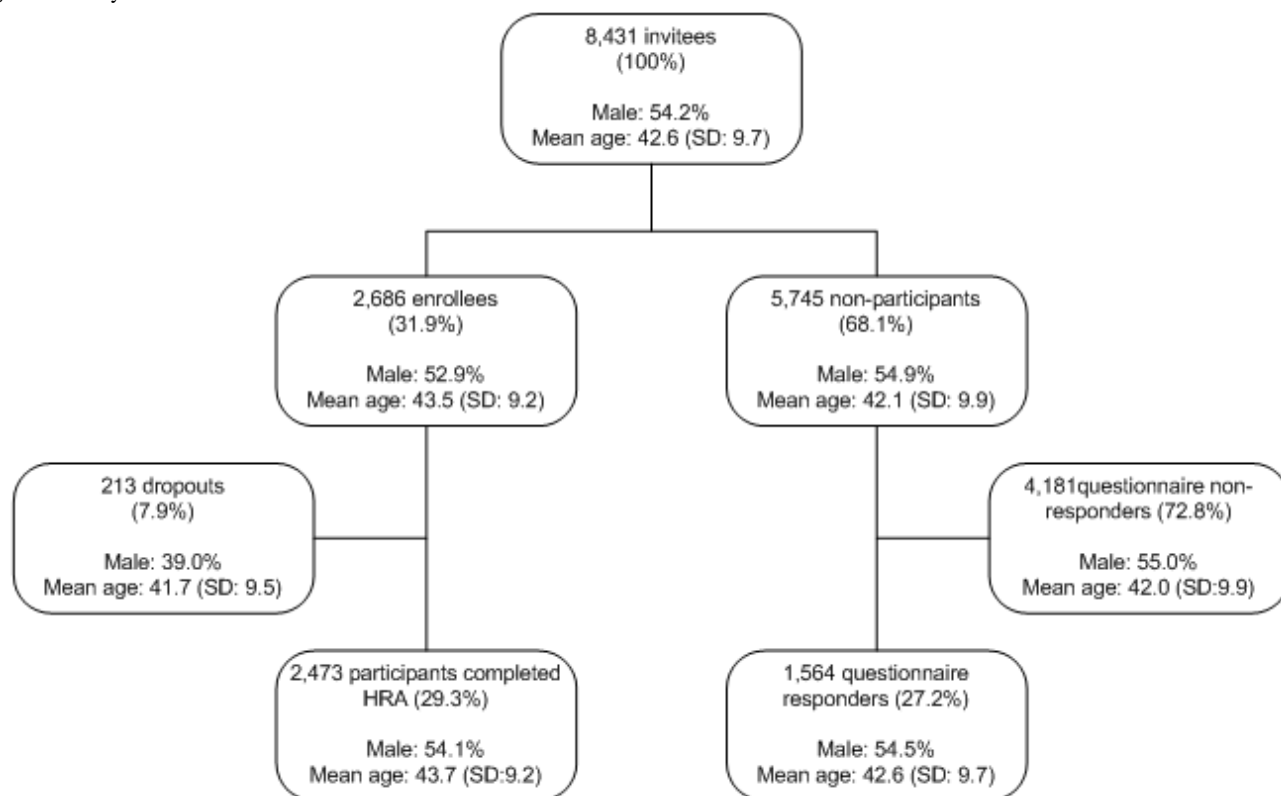
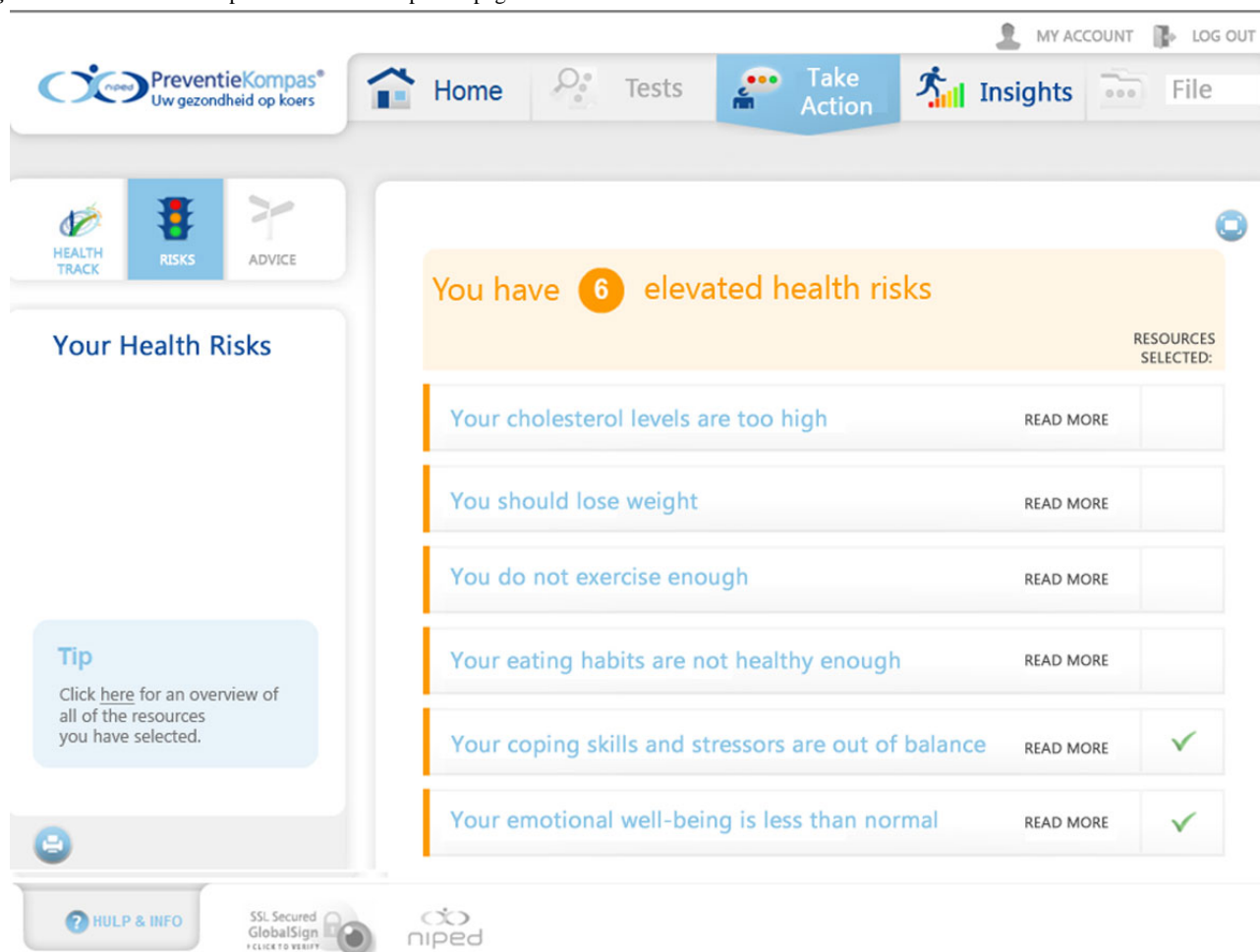


Figure 2. Screenshot of the personal health risk profile page.



Discussion

Principal Results and Comparison With Prior Work

In this study, we evaluated the determinants of participation in a Web-based HRA by comparing participants and nonparticipants with regard to demographics, health behavior, self-rated health, and work-related factors. We found evidence of health-related participation, as workers who were more willing to participate in the HRA engaged in physical activity less frequently, consumed more alcohol, and more frequently experienced some periods of stress at home or work. Nonparticipants rated their overall health less positively and used more tobacco.

Participation in the HRA (31.9%) was similar to the response to the nonparticipant questionnaire (27.2%). The crude analysis pointed towards higher participation among older employees and females. These demographic differences were no longer present in the multivariate analysis. Therefore, the Web-based delivery of the WHPP did not result in selective participation by more highly educated, female, or older employees, which could be explained by the high Internet penetration (94%) in the Netherlands [28]. Although other studies have shown no consistent effect of age on participation [3,15], a meta-analysis performed by Robroek and colleagues (2009) found that women are more likely to participate in WHPPs than men [3]. Also, thus far a number of studies have shown fairly consistently that there is lower participation among employees of lower socioeconomic status [14,15,29-33].

The current study found a strong association between physical activity and HRA participation. The likelihood of participating in the HRA increased as the number of weekdays an employee engaged in physical activity decreased. This result seems to indicate that employees who engage less in physical activity want to know about their state of health, and that those already engaged in frequent physical activity find it less important to participate. However, reports on the influence of physical activity on participation have not been consistent, with some studies pointing towards higher participation in WHPPs among the less physically active [30,34] and other studies indicating higher participation among those with low fitness risk [34] or above-average levels of both habitual activity and physical fitness [35].

Participation in the HRA in our study was also associated with alcohol consumption. Higher weekly alcohol consumption increased the likelihood of participating in the HRA. This finding might be explained by the nonstigmatizing way of addressing alcohol consumption through the Internet. No association between excess alcohol consumption and participation was found in a recent study of a Web-based WHPP [36] or other studies of WHPPs [37].

In the current study, employees who experienced stress at home or at work during the prior year were more likely to participate in the HRA. Two other studies evaluated this association and found similar results [38,39]. These findings suggest that the HRA reaches an important group of workers, as workers under

psychological strain are especially vulnerable to absenteeism and disability [40].

We showed that individuals who rated their health as “moderate” or “bad/very bad” were less likely to participate in the HRA. Self-rated health is associated with physical and mental functioning [18]. In the long run, it is a robust predictor of all-cause mortality and morbidity, and mortality in a range of conditions including cardiovascular disease and cancer [18]. A more immediate association between self-rated health and self-reported absenteeism in the preceding year was found in the current study. Because of these associations, the lack of participation among employees with less-than-positive self-rated health could be interpreted as a general indication that less healthy employees are less likely to participate. One possible reason for this could be that these individuals are currently under treatment for a physical or mental condition. Receiving current medical treatment is an important reason for nonparticipation in WHPPs [38] and was found to be related to nonparticipation in this particular HRA [41]. One could argue that participating in a WHPP is less relevant for those receiving treatment. However, WHPPs and especially broad-based HRAs are designed to screen for a range of chronic diseases and health behaviors, and these programs are likely to benefit individuals who are already receiving medical treatment in other, potentially isolated, areas of health care. Moreover, not everyone with negative self-rated health is receiving medical care. Another reason for lower participation among employees with lower self-rated health could be less healthy employees’ desire to keep their private life and their work life separate. One study found indications that employees with unhealthy lifestyles or who are in poor health are more likely to resist employer interference with employee health [42]. Lower participation among employees with negative self-rated health has been reported in an earlier study on this HRA [41] and other WHPPs [14], but these reports are not consistent [43].

Our study adds to the fairly consistent reports that tobacco users are less likely to participate in WHPPs [30,33,37,38,44]. Most tobacco users are well aware of their habit’s adverse effects and may find they can foresee the outcome and recommendations if they participate in a WHPP. They may find the prospect of such recommendations patronizing and are probably already being confronted with the negative reactions of others in the workplace or at home as a result of their habit. In the HRA under investigation, tobacco users are not encouraged to feel “guilty” or otherwise “pressured” to quit. Intrinsic motivation is recognized as a necessary ingredient for lasting behavior change. Their freedom of choice is affirmed: he or she is respectfully informed of the health benefits of smoking less or quitting and offered resources for bolstering resolve and self-confidence to become smoke-free. However, it is unlikely that the nonjudgmental aspect of this program was communicated to employees prior to their decision of whether or not to participate in the HRA.

This is the second study to evaluate participant characteristics of the HRA, the Prevention Compass. Our study, conducted with a new cohort, addressed two major limitations of the earlier study, which was reported on in 2011 [41]. First, in the 2011 study, only 14% of the nonparticipants completed the online

questionnaire, which formed the basis for the comparison between nonparticipants and participants. As a result, selection bias could have influenced the findings reported in that study. This is hinted at by the substantial difference in reported age between questionnaire responders and nonresponders. Second, we used multivariate analysis in our study. This has the obvious advantage of being able to control for confounding by all other potential determinants. For example, in the 2011 study, it was reported that older employees were likely to participate in the HRA. Also, less self-reported absenteeism was found among participants. We found similar results in the crude analysis of our data. However, in the multivariate analysis, neither age nor absenteeism were still significant determinants. Two of the independent determinants of participation found in the current study—physical activity and alcohol consumption—were not evaluated in the earlier study.

In addition to individual characteristics, program and organizational factors have been linked to participation in WHPPs [37]. Offering financial incentives is one of these factors. Not surprisingly, these incentives increase participation, but one can wonder whether such an external motivator helps to bring about lasting health-behavior change [45]. One of the few studies that investigated the influence of other organizational factors reported a 13% increase in participation in companies with a strong communication strategy [45]. This refers to the extent to which a strategic, comprehensive, integrated communications plan with multiple communications pieces and delivery channels tailored to the employee population is used by companies that offer WHPPs to their work force. Differences in communications strategy during the process of invitation to and inclusion in the HRA could have accounted for some of the variety in participation among the five organizations in the current study. For instance, among the participating organizations in our study, the university medical center had the highest participation (51.7%). In this organization, participation was actively encouraged by upper and middle management, and the HRA was highlighted in the in-house magazine.

By extension, the recruitment message used by organizations can result in selection among participants: whereas Organization A may emphasize one specific feature of the WHPP (eg, “increase your vitality by participating”), Organization B may emphasize another (eg, “screening for health risks”). Following this line of reasoning, the lack of consistent reports in the literature on most individual characteristics of participation may have been caused in part by the widely varying content of recruitment messages. Future research into the reach of WHPPs should consider these and other communication aspects. Based

on the combined insight of individual and organizational characteristics of participation, framing the recruitment message could prove to be an essential tool for companies trying to reach employees with specific risk profiles.

Strengths and Limitations

A limitation of the current study is the low response of the nonparticipants to the nonparticipant questionnaire. Others have been confronted with comparable limitations [36,41]. Individuals who are unwilling to participate in a program are also less likely to respond when asked to participate in a derivative of that program, which in our study was the request to complete a nonparticipant questionnaire. However, in our study, questionnaire responders were of the same age and gender as those who did not respond. Therefore, it is less likely that the reported results have been influenced by selection bias. A strength of the current study is the large size of our study cohort.

No individual characteristics were available for dropouts other than age and gender. This is also a limitation of the current study. Although the number of dropouts (7.9%) was relatively low, their inevitable exclusion from the participant group could have had some influence on the reported findings.

Except for age and gender, which were available from the HRA invitation lists for nearly all (>99.5%) invitees, data on other individual characteristics were collected differently for participants and nonparticipants. For participants, data were collected from the Web-based health questionnaire component of the HRA as part of a larger set of health data collected to generate a personal health report. A separate, short online questionnaire was created to collect data on individual characteristics from the nonparticipants. Some might argue that this divergence in data collection threatens the reliability of the reported findings. However, we estimate this effect to be small, as both participants and nonparticipants completed a set of questions online that were identical with respect to the individual characteristics used in this study.

Conclusion

This study showed health-related participation in a Web-based HRA. With regard to isolated health behaviors (insufficient physical activity, excess alcohol consumption, and stress), those who could benefit most from the HRA were more likely to participate. Employees who rated their overall health as less than positive and tobacco users were less likely to participate. Web-based delivery of the WHPP did not result in selective participation by more highly educated, female, or older employees.

Conflicts of Interest

Coen K Van Kalken and Roderik A Kraaijenhagen are directors and co-owners of the NIPED Research Foundation. This institute developed the program studied. Maurice AJ Niessen is employed full-time as a researcher for the NIPED Research Foundation. All other authors received no additional funding for this study and report no competing interests.

References

1. World Health Organization. *Gaining Health. The European Strategy for the Prevention and Control of Noncommunicable Diseases*. Geneva, Switzerland: WHO Press; 2006.
2. Engbers L. *Monitoring and Evaluation of Worksite Health Promotion Programs Current state of knowledge and implications for practice*. Geneva, Switzerland: WHO Press; 2007.
3. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act* 2009;6:26 [FREE Full text] [doi: [10.1186/1479-5868-6-26](https://doi.org/10.1186/1479-5868-6-26)] [Medline: [19457246](https://pubmed.ncbi.nlm.nih.gov/19457246/)]
4. Colkesen EB, Niessen MA, Peek N, Vosbergen S, Kraaijenhagen RA, van Kalken CK, et al. Initiation of health-behaviour change among employees participating in a web-based health risk assessment with tailored feedback. *J Occup Med Toxicol* 2011;6:5 [FREE Full text] [doi: [10.1186/1745-6673-6-5](https://doi.org/10.1186/1745-6673-6-5)] [Medline: [21388524](https://pubmed.ncbi.nlm.nih.gov/21388524/)]
5. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. *J Med Internet Res* 2008;10(4):e56 [FREE Full text] [doi: [10.2196/jmir.1108](https://doi.org/10.2196/jmir.1108)] [Medline: [19117828](https://pubmed.ncbi.nlm.nih.gov/19117828/)]
6. Franklin PD, Rosenbaum PF, Carey MP, Roizen MF. Using sequential e-mail messages to promote health behaviors: evidence of feasibility and reach in a worksite sample. *J Med Internet Res* 2006;8(1):e3 [FREE Full text] [doi: [10.2196/jmir.8.1.e3](https://doi.org/10.2196/jmir.8.1.e3)] [Medline: [16585028](https://pubmed.ncbi.nlm.nih.gov/16585028/)]
7. Soler RE, Leeks KD, Razi S, Hopkins DP, Griffith M, Aten A, Task Force on Community Preventive Services. A systematic review of selected interventions for worksite health promotion. The assessment of health risks with feedback. *Am J Prev Med* 2010 Feb;38(2 Suppl):S237-S262. [doi: [10.1016/j.amepre.2009.10.030](https://doi.org/10.1016/j.amepre.2009.10.030)] [Medline: [20117610](https://pubmed.ncbi.nlm.nih.gov/20117610/)]
8. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, et al. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health* 2012 Sep;38(5):393-408 [FREE Full text] [doi: [10.5271/sjweh.3275](https://doi.org/10.5271/sjweh.3275)] [Medline: [22270562](https://pubmed.ncbi.nlm.nih.gov/22270562/)]
9. Niessen MA, Kraaijenhagen RA, Dijkgraaf MG, Van Pelt D, Van Kalken CK, Peek N. Impact of a Web-based worksite health promotion program on absenteeism. *J Occup Environ Med* 2012 Apr;54(4):404-408. [doi: [10.1097/JOM.0b013e31824d2e43](https://doi.org/10.1097/JOM.0b013e31824d2e43)] [Medline: [22441493](https://pubmed.ncbi.nlm.nih.gov/22441493/)]
10. Goetzel RZ, Pronk NP. Worksite health promotion how much do we really know about what works? *Am J Prev Med* 2010 Feb;38(2 Suppl):S223-S225 [FREE Full text] [doi: [10.1016/j.amepre.2009.10.032](https://doi.org/10.1016/j.amepre.2009.10.032)] [Medline: [20117606](https://pubmed.ncbi.nlm.nih.gov/20117606/)]
11. Glasgow RE, Klesges LM, Dziewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Educ Res* 2006 Oct;21(5):688-694 [FREE Full text] [doi: [10.1093/her/cyl081](https://doi.org/10.1093/her/cyl081)] [Medline: [16945984](https://pubmed.ncbi.nlm.nih.gov/16945984/)]
12. Bull SS, Gillette C, Glasgow RE, Estabrooks P. Work site health promotion research: to what extent can we generalize the results and what is needed to translate research to practice? *Health Educ Behav* 2003 Oct;30(5):537-549. [Medline: [14582596](https://pubmed.ncbi.nlm.nih.gov/14582596/)]
13. Linnan LA, Emmons KM, Klar N, Fava JL, LaForge RG, Abrams DB. Challenges to improving the impact of worksite cancer prevention programs: comparing reach, enrollment, and attrition using active versus passive recruitment strategies. *Ann Behav Med* 2002;24(2):157-166. [Medline: [12054321](https://pubmed.ncbi.nlm.nih.gov/12054321/)]
14. Conrad P. Who comes to work-site wellness programs? A preliminary review. *J Occup Med* 1987 Apr;29(4):317-320. [Medline: [3585562](https://pubmed.ncbi.nlm.nih.gov/3585562/)]
15. Dobbins TA, Simpson JM, Oldenburg B, Owen N, Harris D. Who comes to a workplace health risk assessment? *Int J Behav Med* 1998;5(4):323-334 [FREE Full text] [doi: [10.1207/s15327558ijbm0504_6](https://doi.org/10.1207/s15327558ijbm0504_6)] [Medline: [16250699](https://pubmed.ncbi.nlm.nih.gov/16250699/)]
16. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999 Sep;89(9):1322-1327 [FREE Full text] [Medline: [10474547](https://pubmed.ncbi.nlm.nih.gov/10474547/)]
17. Van 't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: a longitudinal study. *J Med Internet Res* 2010;12(3):e37 [FREE Full text] [doi: [10.2196/jmir.1345](https://doi.org/10.2196/jmir.1345)] [Medline: [20813716](https://pubmed.ncbi.nlm.nih.gov/20813716/)]
18. Mavaddat N, Kinmonth AL, Sanderson S, Surtees P, Bingham S, Khaw KT. What determines Self-Rated Health (SRH)? A cross-sectional study of SF-36 health domains in the EPIC-Norfolk cohort. *J Epidemiol Community Health* 2011 Sep;65(9):800-806. [doi: [10.1136/jech.2009.090845](https://doi.org/10.1136/jech.2009.090845)] [Medline: [20551149](https://pubmed.ncbi.nlm.nih.gov/20551149/)]
19. Fayers PM, Sprangers MA. Understanding self-rated health. *Lancet* 2002 Jan 19;359(9302):187-188. [doi: [10.1016/S0140-6736\(02\)07466-4](https://doi.org/10.1016/S0140-6736(02)07466-4)] [Medline: [11812551](https://pubmed.ncbi.nlm.nih.gov/11812551/)]
20. 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/)]
21. Rosengren A, Hawken S, Ounpuu S, Sliwa K, Zubaid M, Almahmeed WA, INTERHEART investigators. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364(9438):953-962. [doi: [10.1016/S0140-6736\(04\)17019-0](https://doi.org/10.1016/S0140-6736(04)17019-0)] [Medline: [15364186](https://pubmed.ncbi.nlm.nih.gov/15364186/)]
22. de Zwart BC, Frings-Dresen MH, van Duivenbooden JC. Test-retest reliability of the Work Ability Index questionnaire. *Occup Med (Lond)* 2002 Jun;52(4):177-181 [FREE Full text] [Medline: [12091582](https://pubmed.ncbi.nlm.nih.gov/12091582/)]

23. Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L. The work ability index and single-item question: associations with sick leave, symptoms, and health—a prospective study of women on long-term sick leave. *Scand J Work Environ Health* 2010 Sep;36(5):404-412 [FREE Full text] [Medline: 20372766]
24. Ilmarinen J, Tuomi K, Klockars M. Changes in the work ability of active employees over an 11-year period. *Scand J Work Environ Health* 1997;23 Suppl 1:49-57 [FREE Full text] [Medline: 9247995]
25. Curtin F, Schulz P. Multiple correlations and Bonferroni's correction. *Biol Psychiatry* 1998 Oct 15;44(8):775-777. [Medline: 9798082]
26. Allison PD. Missing Data. In: Sage University Paper Series on Quantitative Applications in the Social Sciences. Thousand Oaks, CA: Sage; 2001:07-136.
27. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: J. Wiley & Sons; 1987.
28. Nederland Europees kampioen internettoegang. URL: <http://www.cbs.nl/nl-NL/menu/themas/bedrijven/publicaties/digitale-economie/artikelen/2012-3636-wm.htm> [accessed 2013-07-08] [WebCite Cache ID 6HxcQSSwC]
29. Glasgow RE, McCaul KD, Fisher KJ. Participation in worksite health promotion: a critique of the literature and recommendations for future practice. *Health Educ Q* 1993;20(3):391-408. [Medline: 8307762]
30. Heaney CA, English P. Are employees who are at risk for cardiovascular disease joining worksite fitness centers? *J Occup Environ Med* 1995 Jun;37(6):718-724. [Medline: 7670919]
31. Brill PA, Kohl HW, Rogers T, Collingwood TR, Sterling CL, Blair SN. The relationship between sociodemographic characteristics and recruitment, retention, and health improvements in a worksite health promotion program. *Am J Health Promot* 1991 Feb;5(3):215-221. [Medline: 10148671]
32. Knight KK, Goetzel RZ, Fielding JE, Eisen M, Jackson GW, Kahr TY, et al. An evaluation of Duke University's LIVE FOR LIFE health promotion program on changes in worker absenteeism. *J Occup Med* 1994 May;36(5):533-536. [Medline: 8027878]
33. Lerman Y, Shemer J. Epidemiologic characteristics of participants and nonparticipants in health-promotion programs. *J Occup Environ Med* 1996 May;38(5):535-538. [Medline: 8733646]
34. Lewis RJ, Huebner WW, Yarborough CM. Characteristics of participants and nonparticipants in worksite health promotion. *Am J Health Promot* 1996 Dec;11(2):99-106. [Medline: 10163602]
35. Shephard RJ, Morgan P, Finucane R, Schimmelfing L. Factors influencing recruitment to an occupational fitness program. *J Occup Med* 1980 Jun;22(6):389-398. [Medline: 7373453]
36. Robroek SJ, Lindeboom DE, Burdorf A. Initial and sustained participation in an internet-delivered long-term worksite health promotion program on physical activity and nutrition. *J Med Internet Res* 2012;14(2):e43 [FREE Full text] [doi: 10.2196/jmir.1788] [Medline: 22390886]
37. Dobbins TA, Simpson JM, Oldenburg B, Owen N, Harris D. Who comes to a workplace health risk assessment? *Int J Behav Med* 1998;5(4):323-334. [doi: 10.1207/s15327558ijbm0504_6] [Medline: 16250699]
38. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Factors associated with non-participation and drop-out in a lifestyle intervention for workers with an elevated risk of cardiovascular disease. *Int J Behav Nutr Phys Act* 2009;6:80 [FREE Full text] [doi: 10.1186/1479-5868-6-80] [Medline: 19951417]
39. Davis KE, Jackson KL, Kronenfeld JJ, Blair SN. Determinants of participation in worksite health promotion activities. *Health Educ Q* 1987;14(2):195-205. [Medline: 3597109]
40. Arends I, van der Klink JJ, Bültmann U. Prevention of recurrent sickness absence among employees with common mental disorders: design of a cluster-randomised controlled trial with cost-benefit and effectiveness evaluation. *BMC Public Health* 2010;10:132 [FREE Full text] [doi: 10.1186/1471-2458-10-132] [Medline: 20230606]
41. Colkesen EB, Kraaijenhagen RA, Frings-Dresen MH, Sluiter JK, van Kalken CK, Tijssen JG, et al. Participation in a workplace web-based health risk assessment program. *Occup Med (Lond)* 2011 Dec;61(8):586-589 [FREE Full text] [doi: 10.1093/occmed/kqr122] [Medline: 21891778]
42. Robroek SJ, van de Vathorst S, Hilhorst MT, Burdorf A. Moral issues in workplace health promotion. *Int Arch Occup Environ Health* 2012 Apr;85(3):327-331 [FREE Full text] [doi: 10.1007/s00420-011-0675-y] [Medline: 21710278]
43. Marshall AL, Leslie ER, Bauman AE, Marcus BH, Owen N. Print versus website physical activity programs: a randomized trial. *Am J Prev Med* 2003 Aug;25(2):88-94. [Medline: 12880874]
44. Baer JT. Improved plasma cholesterol levels in men after a nutrition education program at the worksite. *J Am Diet Assoc* 1993 Jun;93(6):658-663. [Medline: 8389779]
45. 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 Jun;23(5):343-352. [doi: 10.4278/ajhp.08041134] [Medline: 19445438]

Abbreviations

BMI: body mass index

HRA: health risk assessment

WAI: work ability index

WHPP: worksite health promotion program

Edited by G Eysenbach; submitted 09.10.12; peer-reviewed by D Bowen, A Hartzler; comments to author 11.02.13; revised version received 19.03.13; accepted 10.06.13; published 09.08.13.

Please cite as:

*Niessen MAJ, Laan EL, Robroek SJW, Essink-Bot ML, Peek N, Kraaijenhagen RA, Van Kalken CK, Burdorf A
Determinants of Participation in a Web-Based Health Risk Assessment and Consequences for Health Promotion Programs
J Med Internet Res 2013;15(8):e151*

URL: <http://www.jmir.org/2013/8/e151/>

doi: [10.2196/jmir.2387](https://doi.org/10.2196/jmir.2387)

PMID: [23933650](https://pubmed.ncbi.nlm.nih.gov/23933650/)

©Maurice AJ Niessen, Eva L Laan, Suzan JW Robroek, Marie-Louise Essink-Bot, Niels Peek, Roderik A Kraaijenhagen, Coen K Van Kalken, Alex Burdorf. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 09.08.2013. 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

Usage of a Generic Web-Based Self-Management Intervention for Breast Cancer Survivors: Substudy Analysis of the BREATH Trial

Sanne W van den Berg¹, MSc; Esmee J Peters², MSc; J Frank Kraaijeveld³, MSc; Marieke FM Gielissen¹, PhD; Judith B Prins¹, PhD

¹Department of Medical Psychology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

²Department of Medical Oncology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

³IPPZ, ICT & Consultancy for Healthcare, Utrecht, Netherlands

Corresponding Author:

Sanne W van den Berg, MSc

Department of Medical Psychology

Radboud University Nijmegen Medical Centre

PO Box 9101

Nijmegen, 6500 HB

Netherlands

Phone: 31 24 36 13608

Fax: 31 24 36 13425

Email: s.vandenberg@mps.umcn.nl

Abstract

Background: Generic fully automated Web-based self-management interventions are upcoming, for example, for the growing number of breast cancer survivors. It is hypothesized that the use of these interventions is more individualized and that users apply a large amount of self-tailoring. However, technical usage evaluations of these types of interventions are scarce and practical guidelines are lacking.

Objective: To gain insight into meaningful usage parameters to evaluate the use of generic fully automated Web-based interventions by assessing how breast cancer survivors use a generic self-management website. Final aim is to propose practical recommendations for researchers and information and communication technology (ICT) professionals who aim to design and evaluate the use of similar Web-based interventions.

Methods: The BREast cancer ehealth (BREATH) intervention is a generic unguided fully automated website with stepwise weekly access and a fixed 4-month structure containing 104 intervention ingredients (ie, texts, tasks, tests, videos). By monitoring https-server requests, technical usage statistics were recorded for the intervention group of the randomized controlled trial. Observed usage was analyzed by measures of frequency, duration, and activity. Intervention adherence was defined as continuous usage, or the proportion of participants who started using the intervention and continued to log in during all four phases. By comparing observed to minimal intended usage (frequency and activity), different user groups were defined.

Results: Usage statistics for 4 months were collected from 70 breast cancer survivors (mean age 50.9 years). Frequency of logins/person ranged from 0 to 45, total duration/person from 0 to 2324 minutes (38.7 hours), and activity from opening none to all intervention ingredients. 31 participants continued logging in to all four phases resulting in an intervention adherence rate of 44.3% (95% CI 33.2-55.9). Nine nonusers (13%), 30 low users (43%), and 31 high users (44%) were defined. Low and high users differed significantly on frequency ($P<.001$), total duration ($P<.001$), session duration ($P=.009$), and activity ($P<.001$). High users logged in an average of 21 times, had a mean session duration of 33 minutes, and opened on average 91% of all ingredients. Signing the self-help contract ($P<.001$), reporting usefulness of ingredients ($P=.003$), overall satisfaction ($P=.028$), and user friendliness evaluation ($P=.003$) were higher in high users. User groups did not differ on age, education, and baseline distress.

Conclusions: By reporting the usage of a self-management website for breast cancer survivors, the present study gained first insight into the design of usage evaluations of generic fully automated Web-based interventions. It is recommended to (1) incorporate usage statistics that reflect the amount of self-tailoring applied by users, (2) combine technical usage statistics with self-reported usefulness, and (3) use qualitative measures. Also, (4) a pilot usage evaluation should be a fixed step in the development

process of novel Web-based interventions, and (5) it is essential for researchers to gain insight into the rationale of recorded and nonrecorded usage statistics.

Trial Registration: Netherlands Trial Register (NTR): 2935; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2935> (Archived by WebCite at <http://www.webcitation.org/6IkX1ADEV>).

(*J Med Internet Res* 2013;15(8):e170) doi:[10.2196/jmir.2566](https://doi.org/10.2196/jmir.2566)

KEYWORDS

usage evaluation; usage statistics; intervention adherence; user groups; exposure; Internet; Web-based intervention; breast cancer; log file analysis; website use

Introduction

Background

A growing number of women survive breast cancer treatment [1]. The information need is high in these breast cancer survivors [2], and 40-49% of women turn to the Internet for information or support [3-6]. Most breast cancer survivors (70-80%) do not experience severely elevated levels of distress and are not in need of intensive therapist-led psychological treatment [7,8]. Therefore, self-guided Web-based therapeutic interventions [9] seem appropriate to provide easily accessible support to this large number of women at low health care costs. These unguided generic Web-based self-management interventions for breast cancer survivors are emerging and promising [10-12]. However, research data on the use of these type of Web-based interventions are scarce and inconclusive.

Better understanding of website use is an essential step in explaining how Web-based interventions produce behavior change and symptom improvement [13]. The technical usage statistics derived from a website are a representation of the individual processes by which participants use the intervention [14]. These statistics enable us to determine the real-life or observed usage and can be used to calculate adherence rates of Web-based interventions [15]. In addition, the evaluation of usage statistics (usage evaluations or logfile analysis) can reveal important design implications for more effective Web-based interventions [14].

Usage evaluations have been a relatively new area of interest in Internet intervention research. The newly proposed Consolidated Standards of Reporting Trials on eHealth applications (CONSORT-EHEALTH [16]) include the recommendation to report usage parameters. However, practical guidelines are scarce with regard to which usage parameters are preferred to measure observed usage [14,17]. Systematic reviews on the use of Web-based interventions reported a variety of usage statistics, which could be classified into (1) frequency of use (ie, frequency of logins or visits, mean logins during intervention, days on which intervention was visited), (2) duration (ie, length of time logged in), and (3) activity (ie, page views or number of unique pages visited, chapters, or modules completed) [15,18,19]. This multiplicity of usage statistics was also found in usage evaluations of Web-based interventions specifically designed for cancer survivors [10-12]. Deduced from these research findings, at least *frequency*, *duration*, and *activity* should be measured as usage statistics for evaluating the observed usage in Web-based interventions [19].

Evaluating the observed usage is especially important in generic fully automated Web-based interventions. The generic content of these interventions is offered to a heterogeneous group of users, and no professionals are available to tailor the intervention to meet the needs of each individual user. Therefore, we propose the term “self-tailoring” to refer to the degree the user tailors the intervention and selects the content that suits his/her personal situation or needs.

In addition to reporting the observed usage, it is of equal importance to report the intended usage [16,20]. The intended usage is defined prior to evaluation of the observed usage and refers to “the extent to which the developers of the intervention felt that the intervention should be used to achieve the desired effect” [20]. Evaluation of both observed and intended usage can provide insight into whether the intervention was used as envisioned. By comparing the intended usage to the observed usage, a priori defined types of users or user groups can be examined.

Objective

Summarizing, the use of novel generic Web-based interventions is largely unknown and practical guidelines for technical usage evaluations are lacking. Usage evaluations are especially of added value with regard to unguided generic fully automated interventions. It is hypothesized that the use of these generic interventions is more individualized and that users apply a large amount of self-tailoring to the intervention content. Therefore, the present study aims to (1) gain insight into which usage parameters are needed to meaningfully evaluate the usage of generic fully-automated Web-based interventions, by (2) investigating in what amount and how breast cancer survivors use a generic Web-based self-management intervention. Our final aim is to (3) propose practical recommendations for researchers and information and communication technology (ICT) professionals who aim to design and evaluate the use of similar Web-based interventions.

Methods

Participants

This study focused on the analyses of all participants randomly allocated to the intervention group of the BREast cancer ehealth (BREATH) randomized controlled trial (RCT). This two-arm RCT evaluated the efficacy of a Web-based self-management intervention for breast cancer survivors compared to care as usual. Full details of the trial design, eligibility criteria, and patient recruitment have been described

in the study protocol [21]. All participants were (1) women, (2) survivors of primary non-metastatic breast cancer, (3) between 2 and 4 months post treatment, (4) Dutch-speaking, with (5) direct access to a computer with Internet connection, and (6) in possession of an email address.

Intervention

The unguided fully automated Web-based self-management intervention BREATH is based on cognitive behavior therapy (CBT) and aims to foster emotional adjustment after completion of primary curative breast cancer treatment. For a detailed description of the intervention and development process, we refer to the study protocol [21]. In this study, only the details necessary to comprehend the technical usage evaluation are reported. The BREATH intervention covers four phases of recovery after breast cancer, namely Looking Back, Emotional Processing, Strengthening, and Looking Ahead (for a screenshot see [Multimedia Appendix 1](#)). The intervention has a fixed structure with each phase covering 4 weeks. Intervention ingredients (104 in total) include Information (26 scripts), Assignment (48 tasks), Assessment (10 tests), and Video (20 ingredients with thematically clustered video clips extracted from recorded interviews). As a result of the generic character of the intervention, the usage is ad libitum: participants are free to select the intervention ingredients that they find useful or that apply to their personal situation. The first intervention ingredient of the intervention is a self-help contract to stimulate adherence.

The intervention is fully automated following a stepwise weekly access. Each week, on Monday, a reminder email is sent that new information is available. Participants can retrospectively access intervention ingredients of previous weeks but do not have access to forthcoming weeks. In addition to the intervention ingredients, Distress thermometers [22] are available to track the course of experienced distress over the 4-month intervention. Distress thermometers are optional and can be completed with a maximum of 1 per day. Email was used only for reporting technical problems with the website. Based on the session duration of face-to-face CBT, the intended session duration was a maximum of 1 hour per week. The BREATH intervention was developed by a clinical psychologist (JBP) and an eHealth researcher (SWvdb) in close cooperation with ICT professionals (JFK). A multidisciplinary reading committee (including patients, oncology professionals, cancer patient organizations, and patient advocates) reviewed and provided feedback on the thematic content of the intervention [21].

Usage Data Retrieval

The BREATH intervention was developed within the eHealth application myTherapy (IPPZ), designed for online information, communication, and treatment in health care. User-initiated activity in the intervention was determined by monitoring https-server requests. Such requests could be database reads or writes and were logged for various purposes. Database reads could be logged for, for example, logins or opening an intervention ingredient, and database writes for, for example, adding text to an assignment. In most cases, database reads and writes included a timestamp derived from time of the server request. In some cases, timestamps could be combined to

calculate duration. Data were retrieved using logs and database tables used by the Web application myTherapy. Information regarding specific activity (eg, intervention ingredients and internal mail) and user profiles (eg, avatars) was saved or logged to display information to users of myTherapy. For each individual user, user-initiated activity was monitored for a period of 16 weeks.

The occasional absence of data (ie, seconds in log out, precise click path) was in most cases due to design decisions focused on an operational Web application rather than on research purposes. In only a few cases, logging of usage data (eg, logins, login duration) was inaccurate due to rare combinations of, for example, archaic browser type, browser privacy settings, and company network settings. To overcome the problem of patients who forgot to log out, patients were automatically logged out after 30 minutes of inactivity on the website (measured as a 30-minute absence of server requests). During the study period, myTherapy was updated, varying from minor updates (bug fixes) to major updates (minor template changes, improving planning and user interface). In particular, between September 2011 and March 2012, myTherapy suffered irregular short periods of downtime. The total downtime added to less than 1% of the total time.

Outcome Measures: Usage

Frequency, Duration, and Activity

The amount of use of the BREATH intervention was measured with the usage statistics of frequency, duration, and activity. *Frequency* was operationalized as the number of logins per patient during the 4-month period of the intervention. A login was defined as every time a patient signed in to the website for a minimum of 1 minute because no seconds were recorded concerning the logout time. Two types of *duration* were analyzed: session duration and total duration. Session duration was defined as the time (start-stop) of one login in minutes: the time between logging in and logging out. Total duration was the sum of all sessions per patient in minutes. *Activity* was defined as the number of opened intervention ingredients (ie, scripts, tasks, tests, videos) per patient, with a maximum of 104.

To gain insight into how patients used the intervention, we also calculated the distribution (ie, videos, assignments, information, assessments) of the total opened intervention ingredients per patient, using an avatar (yes/no), the number of Distress Thermometers completed, and the number of emails sent to report technical problems with the website, and whether they opened the self-help contract at the beginning of the intervention (yes/no) and signed the contract by filling in their name and the date (considered as actively using the self-help contract). Following the use of 45 assignments and assessments, users were asked whether they perceived these ingredients as useful. This self-reported usefulness was optional to fill in at the end of the assignment or assessment and scored as useful (1), not useful (2), or not filled in (0). For each participant, the proportion of opened ingredients perceived as useful, not useful, or not filled in was calculated.

Last, to evaluate whether the fixed structure was used as such, we calculated how many intervention ingredients were opened

in the phase they were originally planned. To analyze on which days participants log in to the intervention during the week, each login was coded into a nominal variable representing Sunday-Saturday (1-7). Results on how the intervention was used were reported for low and high users.

Intervention Adherence and Nonusage Attrition

In this study, intervention adherence solely referred to the extent to which participants were exposed to the content of the intervention, not adherence to research protocol assessments (eg, filling out questionnaires) [20,23,24]. In addition, nonusage attrition [24] (or nonadherence) referred to the proportion of participants who stopped using the intervention over time. In Internet intervention research, there is a lack of agreement about which definitions and usage statistics should best be used to measure adherence or nonusage attrition [15]. In the current study, intervention adherence was defined as user persistence or *continuous usage*: the proportion of patients who started using the intervention and continued to log in (at least once) during all four phases. Nonusage attrition was defined as *intermittent usage*: the proportion of patients who did not log in during all four phases of the intervention. Continuous and intermittent usage were measured based on frequency of logins. For each participant, it was calculated in which phases (1-4) and weeks (1-16) logins took place.

User Groups

To evaluate how participants used the intervention differently, user groups were calculated by comparing the intended usage to the observed usage. The minimal intended frequency of logins as formulated by the developers of the BREATH intervention was a minimum of 10 times over the course of the intervention and was based on the frequency of face-to-face CBT. Also, the intervention ingredients of 1 week should take a maximum of 1 hour to complete. The minimal intended activity was opening a minimum of 50% of the total 104 intervention ingredients, because not all ingredients of the generic intervention will apply to the personal situation of every user. Table 1 gives an overview of the classification of four user groups based on minimal intended frequency and activity. To calculate user groups, the observed frequency and activity was cross-tabulated within a 4x4 matrix of intended frequency and activity.

Outcome Measures: Other

Baseline Survey

At baseline, before randomization, participants of the BREATH RCT filled in an online survey with questions concerning sociodemographic characteristics (eg, age, marital status, children, education, employment status), medical characteristics (eg, type of adjuvant therapy, use of hormonal therapy), and psychological questionnaires (for a full overview see [21]). Education was measured using a 7-point scale [25] ranging from primary education not finished (1) to master's degree (7). For this study, with regard to psychological questionnaires, only the Hospital Anxiety and Depression Scales (HADS) [26] was reported to assess baseline general distress [27]. The total score of the HADS (HADS-T) has demonstrated good reliability and validity in oncology patients [28,29]. A HADS-T of ≥ 11 represented elevated levels of distress indicative for mental disorders [30].

Evaluation Survey

After the intervention (4 months after baseline), participants completed an online survey including an evaluation of the intervention. For the intervention evaluation, two single-item measures were examined: overall satisfaction ("Which grade would you give to the overall intervention?") and user friendliness ("Which grade would you give to the user friendliness of the intervention?"). These measures were scored on a 10-point scale ranging from 1 (very bad) to 10 (very good). For qualitative results, participants were asked to report points for improvements of the intervention.

Statistical Analyses

All analyses were performed using SPSS 20. For all nondescriptive outcome measures, the amount, the percentage, and the Wilson confidence interval (CI) were reported. Usage statistics, sociodemographic, and medical characteristics were not normally distributed as indicated with the Kolmogorov-Smirnov test (<0.05) and therefore were first analyzed using nonparametric tests. To facilitate interpretation, parametric tests were reported, since results did not differ from nonparametric tests. Pearson's correlation coefficients were reported between technical usage statistics (ie, frequency, session duration, total duration, and activity) and between usage statistics and the patient characteristics. To assess differences between user groups, *t* tests, Pearson's chi-square tests, and Fisher exact tests were conducted. A two-sided $\alpha=0.05$ level of significance was used for all analyses.

Table 1. Classification of user groups based on minimal intended frequency and activity.

	Minimal intended frequency	Minimal intended activity
Nonusers	0	0
Low users	1	1%
Intended users	10	50%
High users	17	75%

Results

Sociodemographic and Medical Characteristics

Seventy participants were included in the study sample and had been in the position to log in to the BREATH intervention for a period of 4 months. Usage statistics were recorded from November 2010 until August 2012. Of all participants, mean age was 50.9 (SD 8.31), the mean education level on a 7-point scale was 5 (SD 1.63), and 1 participant did not have Dutch nationality. Forty percent of the patients were employed (28/70), 37% (33/70) received full or partial disablement insurance or were on sick leave, 83% of the participants (58/70) were married or living together with a partner, and 87% of the participants (61/70) had children. All participants were treated with surgery and adjuvant therapy for breast cancer: 27% (19/70) received only chemotherapy, 4% (3/70) received only radiotherapy, and 69% (48/70) received both chemotherapy and radiotherapy. In addition, 66% of the participants (46/70) received hormonal therapy during the intervention period. At baseline, 27% of the participants (19/70) experienced elevated levels of distress based on HADS-T \geq 11.

Frequency, Duration, and Activity

Participants demonstrated a large variability in intervention usage over the 4 months in which the intervention was available. Frequency ranged from 0 to 45 logins (mean 11, SD 7), and 10% (7/70) of the participants never logged in to the intervention. Total duration per participant ranged from 0 to 2324 minutes (38.7 hours), with a mean total duration per participant of 337.2 minutes (SD 163.7), which equals 5.6 hours. The mean of the average session duration per patient was 24.7 minutes (SD 16.1). Activity ranged from opening none to all intervention ingredients, with a mean of opened intervention ingredients per participant of 49.9 (SD 42.8), and 13% (9/70)

of the participants never opened an intervention ingredient. Frequency was positively correlated with total duration ($r=.83$), session duration ($r=.40$), and activity ($r=.84$), and high activity was associated with a longer total ($r=.75$) and session duration ($r=.55$). All correlations were significant on the $P<.001$ level. Correlations between total and session duration were not calculated because total duration was calculated with session duration.

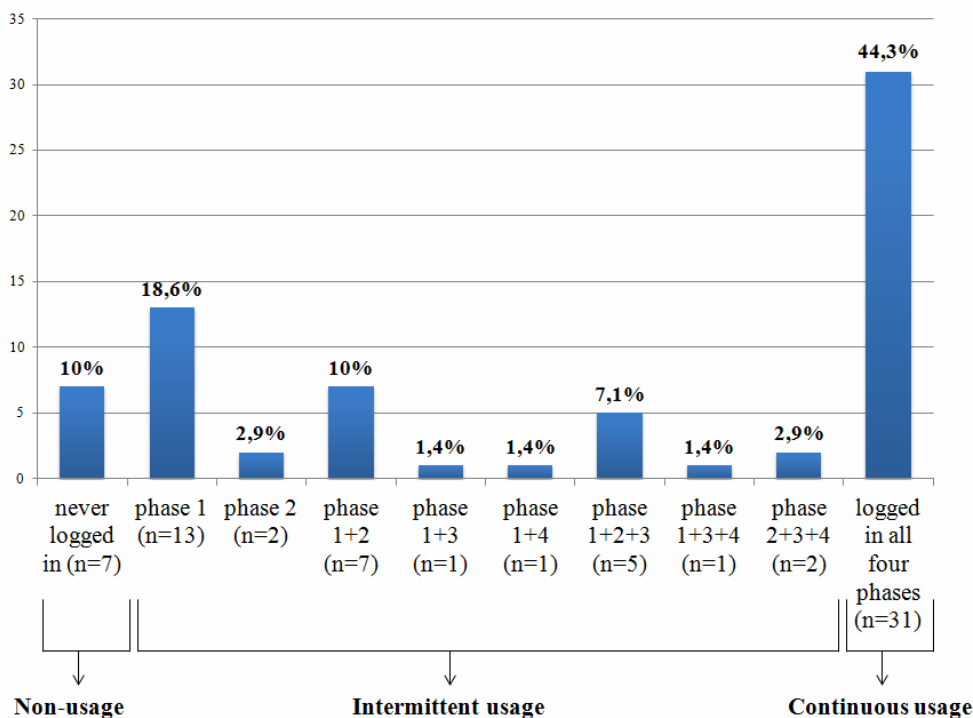
With regard to how patients used the intervention, 69% of the participants (48/70; 95% CI 56.97-78.24) opened the self-help contract, and 17% (12/70; 95% CI 10.09-27.62) made use of an avatar. Of all participants, 63% (44/70) filled in at least one Distress Thermometers: median 2 and a maximum of 13. Seven participants sent emails to the researcher concerning technical problems with the intervention. There were significant differences between the login days ($P<.001$), with 28% (CI 24.80-31.17) of all logins ($n=757$) being on the day the weekly reminder was sent (Monday).

Intervention Adherence and Nonusage Attrition

Figure 1 shows the intervention adherence (defined as continuous usage) and nonusage attrition (defined as intermittent usage) based on logins during the four intervention phases. Of the total sample, 31 participants logged in to the intervention website during all four phases, resulting in a continuous usage of 44.3% (95% CI 33.2-55.9). Of these participants, only 6 logged in during all 16 weeks of the intervention.

Seven participants (10%) never logged in to the website and were thus never exposed to the intervention content. Intermittent usage was 45.7% (32/70): 13 participants (18.6%) only logged in during the first phase, and 2 participants (2.9%) only logged in during the second phase. Nine users (12.6%) logged in during two of the four phases, and 8 users (11.4%) logged in only during three phases.

Figure 1. Continuous usage and intermittent usage based on logins during 4 intervention phases (n=70).



User Groups

This study showed three user groups based on the comparison of intended versus observed frequency and activity: 9 nonusers, 30 low users, and 31 high users. Seven nonusers never logged in, and 2 nonusers logged in once but did not open any intervention ingredients. Only 2 participants met the classification of intended user as specified beforehand. Not being considered a substantial group, these 2 intended users were listed as high users. Low and high users differed significantly on frequency ($P<.001$), total duration ($P<.001$), session duration ($P=.009$), and activity ($P<.001$). Low users logged in with an average of 3.6 times (SD 2.6) over the course of the 4-month intervention and had a mean session duration of 23.5 minutes (SD 12.3). The mean total duration that low users spent on the website was 81.1 minutes (SD 75.5) in which they opened a mean of 18.8/104 ingredients (SD 17.2). High users logged in with an average of 21 times (SD 9), which is more than once a week, and had a mean session duration of 32.8 minutes (SD 14.4). The mean total duration that high users spent on the website was 682.7 minutes (SD 443 minutes), which equals 11 hours and 22 minutes. During this time, high users opened on average 91% of all intervention ingredients (mean 94.5/104 ingredients, SD 12.8).

Group characteristics of the three user groups are reported in Table 2. On baseline distress, sociodemographic, and medical characteristics, no significant differences were found between nonusers versus users (low and high users), and low users versus high users.

With regard to how the intervention was used, high users completed significantly more Distress Thermometers (mean 5, SD 2.5) compared to low users (mean 1, SD 1.5; $P<.001$). In addition, all high users (100%; 31/31) opened the self-help contract at the beginning of the intervention, versus 57% (17/30)

of the low users ($P<.001$). Following the opening of the self-help contract, 84% (26/31) of the high users also signed the contract versus 53% (9/17) of the low users ($P<.001$).

Self-reported usefulness was gathered for the majority of the intervention ingredients that required active input from users (assignments and assessments). The proportion of opened ingredients perceived as useful was higher in high users (mean 67%, SD 21%) compared to low users (mean 44%, SD 25%; $P<.001$). High users filled in the self-reported usefulness significantly more often than low users (mean proportion not filled 16%, SD 17%, versus mean 36%, SD 29%; $P=.003$). The proportion of opened ingredients reported as not useful was low and did not differ between high users (mean 18%, SD 18%) and low users (mean 21%, SD 20%; $P=.557$).

With regard to following the fixed structure, low users opened 19.7% of the intervention ingredients in a later phase than the ingredients were planned. High users followed the structure more and opened only 5.7% of the intervention ingredients in a later phase. The standard intervention distribution of the 104 ingredients was 46% assignments, 25% information, 19% videos, and 10% assessments. Figure 2 displays the distribution of intervention ingredients for each participant. Both low and high users did not show a strong preference in the type of opened intervention ingredients, for example, opening only videos. The proportion of opened assignments (40% vs 45%; $P=.178$), information (26% vs 25%; $P=.850$), and videos (21% vs 20%; $P=.653$) did not differ between low and high users. Low users opened proportionally more assessments compared to the high users (15% vs 10%; $P=.036$). However, this was related to the fact that all assessments were in the first two phases and low users opened predominantly ingredients in these first phases of the intervention. Last, high and low users did not differ on using an avatar or sending emails to the researcher about technical problems.

Figure 2. Distribution of total opened intervention ingredients per participants (n=70).

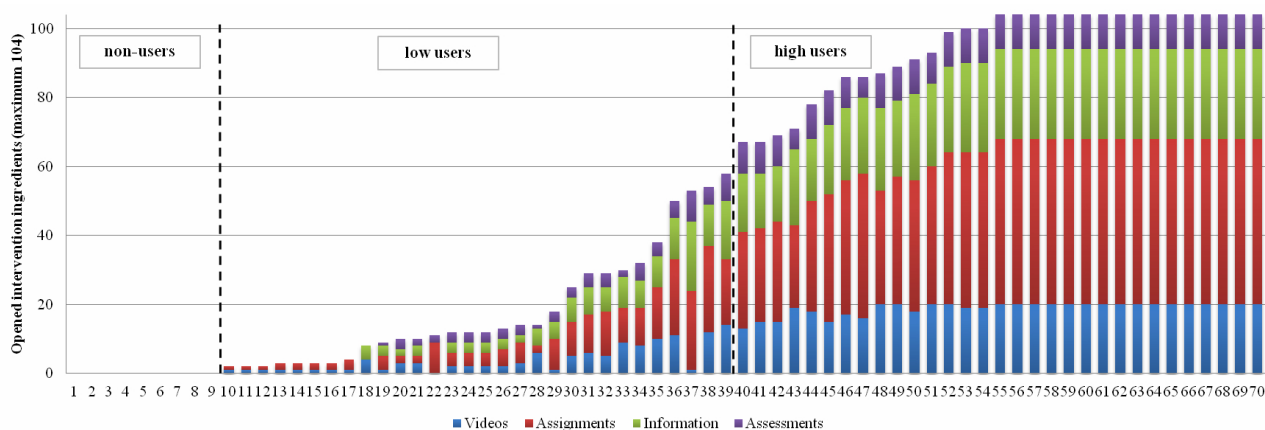


Table 2. Group characteristics (sociodemographic, social, medical) and baseline distress of the three user groups (n=70).

Characteristics	Nonusers (n=9)	Users (n=61)		P value
		Low users (n=30)	High users (n=31)	Low vs high users
Age, mean (SD)	51.9 (7)	51.83 (8.73)	49.7 (8.3)	.33 ^b
Education (1-7), mean (SD)	5 (1.3)	4.9 (1.5)	5.2 (1.3)	.36 ^b
Married/cohabiting, n (%)	6 (66.7)	27 (90)	25 (80.6)	.47 ^c
Children, n (%)	8 (88.9)	28 (93.3)	25 (80.6)	.26 ^c
Employment				
Paid job, n (%)	2 (22.2)	14 (46.7)	12 (38.7)	.53 ^d
Disablement insurance act or sick leave, n (%)	5 (55.6)	10 (33.3)	18 (58.1)	.053 ^d
Adjuvant treatment				
Both chemotherapy and radiotherapy, n (%)	6 (66.7)	22 (73.3)	20 (64.5)	.46 ^d
Hormonal therapy, n (%)	8 (88.9)	16 (53.3)	22 (71)	.16 ^d
Baseline distress				
HADS-T ^a baseline, mean (SD)	10.6 (11.1)	8.6 (5.8)	9.1 (6.1)	.75 ^b
HADS-T ^a ≥ 11, n (%)	4 (44.4)	9 (30)	10 (32.3)	.85 ^d

^aHADS-T=Hospital Anxiety and Depression Scale-total score.

^bIndependent samples *t* test.

^cFisher's exact test.

^dPearson chi-Square.

Evaluation Survey

Post-intervention evaluation surveys were filled in by 20 low users and 30 high users. Two nonusers erroneously filled in evaluation surveys, but since they were never exposed to the intervention, these were left out of the analyses. High users evaluated both overall satisfaction (mean 7, SD 1.20 vs mean 5.75, SD 2.20; $P=.028$) and user friendliness (mean 7.27, SD 1.34 vs mean 5.58, SD 1.18; $P=.003$) significantly higher than low users. Twenty-five participants (11 low users and 14 high users) actively stated points for improvements to the intervention. The top three points for improvements were (1) possibility to get access to the intervention sooner after completion of breast cancer treatment (6/25, 24%), (2) lack of practical information (eg, on prostheses, wigs, bras; 4/25, 16%), and (3) poor user friendliness of logging in (security code sent to mobile phone; 3/25, 12%).

Discussion

Summary

The current formative usage evaluation of a self-management website for breast cancer survivors illustrated the supposed diverse and individualized usage of generic fully automated Web-based interventions. Evaluation of only the amount of usage on group level did not provide a valuable representation of the real-life exposure to the generic self-management intervention. Usage data on how the intervention was used proved to be informative and revealed that 44.3% of the women

continued using the BREATH-intervention over the 4-month period. Also, the comparison of intended versus observed usage showed three different user groups. A small proportion of participants were never, or only once, exposed to the intervention and were classified as nonusers. While the intended user group proved to be nonsubstantial, two equally large groups of active users were defined: low users and high users. Apart from the significant differences in usage statistics, low and high users were found to have a distinctive way of how they used the intervention. High users had a more homogeneous and consistent usage compared to low users. High users exceeded the intended frequency and activity, signed the self-help contract at the beginning of the intervention, and followed the fixed time-locked structure of the intervention. Although technical usage statistics did not provide information on the amount of self-tailoring users applied after they opened intervention content, data on self-reported usefulness showed that high users perceived the majority of opened intervention ingredients as useful. User groups did not differ in pre-intervention distress, sociodemographic, or medical characteristics.

The choice, or technical availability, of usage statistics plays a crucial part in usage evaluations and poses hazards to misinterpretations. For example in this study, solely based on the finding that high users opened almost all intervention ingredients could lead to the premature conclusion that all ingredients were useful to these participants. The fact that on group level, no preferences were found in opening intervention ingredients could add to this misinterpretation. However, based on technical usage data, it was not possible to conclude that

high users valued all ingredients equally, since data on re-opening intervention ingredients were lacking. Data on self-reported usefulness provided this missing information and proved to be essential in making conclusions about how users self-tailored the content of this generic self-management intervention. In contrast to low users, high users consistently reported about the usefulness of intervention ingredients, and they perceived the majority of opened intervention content useful. Therefore, we concluded that high users actively used the full intervention content.

We also concluded that high users self-tailor logins to their own timetable instead of logging in during each intervention week. This was based on the finding that only 6 high users logged in during all 16 weeks. This sheds new light on the mean frequency of logins of 21 times of the high users. Apparently, high users do not log in during some weeks but catch up during the next week by returning to the website several times that week. Combined with the knowledge that high users follow the phase-structure when it comes to opening intervention ingredients in the planned phase, this might imply that time-locks can be broader in the future. Session duration was around 30 minutes in both low and high users and was lower than the maximum intended session duration of 1 hour, which implies that the natural session duration of the BREATH-intervention is half an hour.

Intervention Adherence

As a result of the lack of agreement about how best to define and measure adherence, we have chosen to define intervention adherence as continuous usage based on frequency of logins. In order to be transparent, we consistently reported “continuous usage” throughout the current manuscript or provided our operationalization in addition to adherence: “intervention adherence (continuous usage)”. In a systematic review, Donkin et al [15] found that most studies on Web-based interventions reported adherence based on frequency of logins. However, it is recommended to use a composite measure encompassing a variety of usage statistics for the calculation of adherence [15]. High correlations found between frequency, total duration, and activity in the present study suggest that these three usage statistics measure a similar construct of continuous usage and are therefore interchangeable in analyses of adherence in this study. Whether they are also interchangeable in the analyses of effectiveness needs further research, since Donkin et al [15] found that activity (defined as completion of modules) was most consistently related to outcomes in psychological health interventions. Confirmed by other studies [31], in the current usage evaluation duration was found to be the least precise and therefore least reliable usage statistic. Since it is unknown what users do when a website is opened on their computer screen, time spent on a website provides the least reliable estimation of exposure to an intervention content.

Information on both adherence and nonusage attrition can be similarly informative in future evaluations of effect. Previous research has demonstrated that nonadherers can benefit equally as adherers from the intervention content they completed [32]. In the current study, it is possible that the participants who logged in continuously or intermittently during three out of four

phases, experienced an early effect, which made further use of the intervention redundant. Different factors may predict intervention adherence in Web-based interventions [13], such as support provided by a therapist or coach [18,33], intervention characteristics, being studied in the context of a RCT design, a high frequent intended usage, and the use of persuasive technology [20]. Sending email reminders is part of persuasive technology [34]. The positive influence of sending weekly email reminders on intervention adherence (in the current study defined as continuous usage) was confirmed by the fact that 28% of all logins were on the same day the email reminder was sent. Email reminders were standard, but every month the reminder contained a preview of the intervention content of the upcoming 4 weeks, which might also have had a beneficial effect on revisiting the intervention [35].

Predictors of Usage

In this study, user groups only differed in usage statistics, which is how they were classified. With regard to how the intervention was used, high users signed the self-help contract more often and reported more consistently on the usefulness of ingredients compared to low users. However, in the current study we lacked data to know the causality of these findings. At this moment, we do not know whether signing the self-help contract and reporting usefulness are predictors of high use, or whether high use predicts signing the self-help contract and reporting usefulness. More research is needed to determine whether and how intervention characteristics (such as a self-help contract) or user characteristics (such as motivation, positive expectations) can influence high usage.

In addition, no specific sociodemographic, medical or personal characteristics were found that distinguished between user groups, supporting our hypothesis that the present generic fully automated intervention could be acceptable for a broad range patients. However, this also led to a lot of unanswered questions about possible predictors of usage. It is possible that other characteristics not taken into consideration in the present study predict who is going to be a low or high user. For example, information on pre-intervention needs was lacking. Although distress was not related to the observed usage, distress screening does not uncover unmet needs in posttreatment cancer survivors [36]. Other possible predictors of usage could be computer experience, social support, or illness burden. In a Web-based illness management support system for breast and prostate cancer patients (WebChoice), the level of computer experience proved to be a predictor of use, whereas low social support and high illness burden were associated with high use of specific intervention components [37]. Another explanation for the absence of predictors could be that the usage behavior itself predicts whether users continue to use the intervention or do not log in again.

Pitfalls and Limitations

The most important pitfall of the current study was the absence of usage data on re-visiting or re-opening intervention ingredients due to design decisions focused on the intervention website being operational. As a result, we lacked technical usage information on patient preferences of certain types of intervention ingredients after their first opening. Data on

self-reported usefulness provided nontechnical data on this matter and allowed us to make some statements about self-tailoring.

The current study also lacked essential qualitative knowledge about reasons to stop or continue using the intervention. For example, in this study overall satisfaction and user friendliness evaluation of the intervention was higher in high users, but the causality of this finding needs further qualitative investigation. Stopping with the intervention might be negatively related to characteristics of the website (eg, user friendliness, appearance), the content of the intervention (eg, did not meet the patients needs), or the patient (eg, too burdensome, concurrent life events).

Recommendations for Researchers and ICT Professionals

Based on the pitfalls encountered in the current study we formulated the following recommendations for researchers and ICT-professionals conducting usage evaluations of generic fully automated Web-based interventions. First, choose usage statistics that give insight into the amount of self-tailoring that participants apply to the intervention content and structure. This implies to record both singular usage statistics (frequency, duration, activity) and composite usage statistics (time spent per ingredient, click-patterns, re-opening, or span of use [14]). Second, combine technical usage statistics with self-reported usefulness to gain additional information on specific intervention components. The question of whether an intervention component is useful or not is easily implemented at the end of each component and takes little effort for participants. In case of missing technical data, self-reported usefulness can provide valuable insight in the amount of self-tailoring applied by users.

Third, combine technical usage statistics with qualitative measures (such as semistructured telephone interviews or online focus groups) for a comprehensive usage evaluation. Fourth, conduct a pilot usage evaluation with a variety of usage statistics as a fixed step in the iterative development process of Internet interventions. This way, decisions can be made about which usage statistics should meaningfully be taken into account, or left out, in the final evaluation of usage. Last, gain insight into the rationale of recorded and nonrecorded usage statistics. Researchers with basic knowledge of ICT combined with ICT professionals with basic knowledge about conducting research facilitate effective communication and clear agreements about usage evaluations.

Conclusion

This study underscores the added value of evaluating usage statistics of generic Web-based interventions as a realistic estimation of exposure to intervention content. To the best of our knowledge, the present study gained first insight into the design of technical usage evaluations of generic fully automated Web-based interventions. Overall, and in concordance with research on more interactive eHealth applications [38], results suggest that investigating how generic fully automated Web-based interventions are used is far more informative than the amount of exposure. Usage statistics should be chosen accordingly. Further, it is recommended to collect both singular and composite usage statistics, include self-reported usefulness, and to pilot test a variety of usage statistics to aid decision making of meaningful usage parameters. Last, shared knowledge about ICT and conducting research is helpful in developing a meaningful rationale of technically recorded usage statistics of generic Web-based interventions.

Acknowledgments

The BREATH study is funded by Pink Ribbon, the Netherlands.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of the Web-based non-guided self-management intervention BREATH (in Dutch).

[[PNG File, 560KB - jmir_v15i8e170_app1.png](#)]

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011;61(2):69-90 [[FREE Full text](#)] [doi: [10.3322/caac.20107](https://doi.org/10.3322/caac.20107)] [Medline: [21296855](https://pubmed.ncbi.nlm.nih.gov/21296855/)]
2. Cappiello M, Cunningham RS, Knobf MT, Erdos D. Breast cancer survivors: information and support after treatment. *Clin Nurs Res* 2007 Nov;16(4):278-93; discussion 294. [doi: [10.1177/1054773807306553](https://doi.org/10.1177/1054773807306553)] [Medline: [17991908](https://pubmed.ncbi.nlm.nih.gov/17991908/)]
3. Afshari M, Ghani MFA, Radzi NM. Use of the internet by breast cancer patients. *Sci Res Essays* 2011;6(17):1992-2248.
4. Eysenbach G. The impact of the Internet on cancer outcomes. *CA Cancer J Clin* 2003;53(6):356-371. [Medline: [15224975](https://pubmed.ncbi.nlm.nih.gov/15224975/)]
5. Fogel J, Albert SM, Schnabel F, Ditkoff BA, Neugut AI. Use of the Internet by women with breast cancer. *J Med Internet Res* 2002;4(2):E9 [[FREE Full text](#)] [doi: [10.2196/jmir.4.2.e9](https://doi.org/10.2196/jmir.4.2.e9)] [Medline: [12554556](https://pubmed.ncbi.nlm.nih.gov/12554556/)]
6. Pereira JL, Koski S, Hanson J, Bruera ED, Mackey JR. Internet usage among women with breast cancer: an exploratory study. *Clin Breast Cancer* 2000 Jul;1(2):148-53; discussion 154. [doi: [10.3816/CBC.2000.n.013](https://doi.org/10.3816/CBC.2000.n.013)] [Medline: [11899653](https://pubmed.ncbi.nlm.nih.gov/11899653/)]

7. Helgeson VS, Snyder P, Seltman H. Psychological and physical adjustment to breast cancer over 4 years: identifying distinct trajectories of change. *Health Psychol* 2004 Jan;23(1):3-15. [doi: [10.1037/0278-6133.23.1.3](https://doi.org/10.1037/0278-6133.23.1.3)] [Medline: [14756598](https://pubmed.ncbi.nlm.nih.gov/14756598/)]
8. Henselmans I, Helgeson VS, Seltman H, de Vries J, Sanderman R, Ranchor AV. Identification and prediction of distress trajectories in the first year after a breast cancer diagnosis. *Health Psychol* 2010 Mar;29(2):160-168. [doi: [10.1037/a0017806](https://doi.org/10.1037/a0017806)] [Medline: [20230089](https://pubmed.ncbi.nlm.nih.gov/20230089/)]
9. 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/)]
10. Beatty L, Koczwara B, Wade T. 'Cancer Coping Online': A pilot trial of a self-guided CBT internet intervention for cancer-related distress. *EJAP* 2011 Dec 20;7(2):17. [doi: [10.7790/ejap.v7i2.256](https://doi.org/10.7790/ejap.v7i2.256)]
11. Carpenter KM, Stoner SA, Schmitz K, McGregor BA, Doorenbos AZ. An online stress management workbook for breast cancer. *J Behav Med* 2012 Dec 2:-. [doi: [10.1007/s10865-012-9481-6](https://doi.org/10.1007/s10865-012-9481-6)] [Medline: [23212928](https://pubmed.ncbi.nlm.nih.gov/23212928/)]
12. Duffecy J, Sanford S, Wagner L, Begale M, Nawacki E, Mohr DC. Project onward: an innovative e-health intervention for cancer survivors. *Psychooncology* 2013 Apr;22(4):947-951. [doi: [10.1002/pon.3075](https://doi.org/10.1002/pon.3075)] [Medline: [22438297](https://pubmed.ncbi.nlm.nih.gov/22438297/)]
13. 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/)]
14. Han JY. Transaction logfile analysis in health communication research: challenges and opportunities. *Patient Educ Couns* 2011 Mar;82(3):307-312. [doi: [10.1016/j.pec.2010.12.018](https://doi.org/10.1016/j.pec.2010.12.018)] [Medline: [21277146](https://pubmed.ncbi.nlm.nih.gov/21277146/)]
15. 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/)]
16. 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/)]
17. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. *J Med Internet Res* 2006;8(3):e15 [FREE Full text] [doi: [10.2196/jmir.8.3.e15](https://doi.org/10.2196/jmir.8.3.e15)] [Medline: [16954125](https://pubmed.ncbi.nlm.nih.gov/16954125/)]
18. 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;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/)]
19. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. *Health Educ Behav* 2011 Feb;38(1):49-62. [doi: [10.1177/1090198110372878](https://doi.org/10.1177/1090198110372878)] [Medline: [21189422](https://pubmed.ncbi.nlm.nih.gov/21189422/)]
20. 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/)]
21. van den Berg SW, Gielissen MF, Ottevanger PB, Prins JB. Rationale of the BREast cancer e-healTH [BREATH] multicentre randomised controlled trial: an internet-based self-management intervention to foster adjustment after curative breast cancer by decreasing distress and increasing empowerment. *BMC Cancer* 2012;12:394 [FREE Full text] [doi: [10.1186/1471-2407-12-394](https://doi.org/10.1186/1471-2407-12-394)] [Medline: [22958799](https://pubmed.ncbi.nlm.nih.gov/22958799/)]
22. Tuinman MA, Gazendam-Donofrio SM, Hoekstra-Weebers JE. Screening and referral for psychosocial distress in oncologic practice: use of the Distress Thermometer. *Cancer* 2008 Aug 15;113(4):870-878 [FREE Full text] [doi: [10.1002/cncr.23622](https://doi.org/10.1002/cncr.23622)] [Medline: [18618581](https://pubmed.ncbi.nlm.nih.gov/18618581/)]
23. 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/)]
24. 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/)]
25. Verhage F. Intelligence and Age in a Dutch Sample. *Human Development* 1965;8(4):238-245. [doi: [10.1159/000270308](https://doi.org/10.1159/000270308)]
26. 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/)]
27. Norton S, Cosco T, Doyle F, Done J, Sacker A. The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis. *J Psychosom Res* 2013 Jan;74(1):74-81. [doi: [10.1016/j.jpsychores.2012.10.010](https://doi.org/10.1016/j.jpsychores.2012.10.010)] [Medline: [23272992](https://pubmed.ncbi.nlm.nih.gov/23272992/)]
28. Alexander S, Palmer C, Stone PC. Evaluation of screening instruments for depression and anxiety in breast cancer survivors. *Breast Cancer Res Treat* 2010 Jul;122(2):573-578. [doi: [10.1007/s10549-009-0669-6](https://doi.org/10.1007/s10549-009-0669-6)] [Medline: [19960243](https://pubmed.ncbi.nlm.nih.gov/19960243/)]
29. Annunziata MA, Muzzatti B, Altoè G. Defining hospital anxiety and depression scale (HADS) structure by confirmatory factor analysis: a contribution to validation for oncological settings. *Ann Oncol* 2011 Oct;22(10):2330-2333 [FREE Full text] [doi: [10.1093/annonc/mdq750](https://doi.org/10.1093/annonc/mdq750)] [Medline: [21339383](https://pubmed.ncbi.nlm.nih.gov/21339383/)]

30. Vodermaier A, Millman RD. Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: a systematic review and meta-analysis. *Support Care Cancer* 2011 Dec;19(12):1899-1908. [doi: [10.1007/s00520-011-1251-4](https://doi.org/10.1007/s00520-011-1251-4)] [Medline: [21898134](https://pubmed.ncbi.nlm.nih.gov/21898134/)]
31. Crutzen R. Utilising exposure measures of internet-delivered interventions. In: Olsen E, editor. *Behaviour change & technology*. Oslo: Changetech; 2009.
32. Hilvert-Bruce Z, Rossouw PJ, Wong N, Sunderland M, Andrews G. Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behav Res Ther* 2012 Aug;50(7-8):463-468. [doi: [10.1016/j.brat.2012.04.001](https://doi.org/10.1016/j.brat.2012.04.001)] [Medline: [22659155](https://pubmed.ncbi.nlm.nih.gov/22659155/)]
33. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res* 2011;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
34. Oinas-Kukkonen HH, Arjuma M. Persuasive systems design: Key issues, process model, and system features. *Commun Assoc Inf Syst* 2009;24(1):485-500.
35. 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/)]
36. van Scheppingen C, Schroevens MJ, Smink A, van der Linden YM, Mul VE, Langendijk JA, et al. Does screening for distress efficiently uncover unmet needs in cancer patients? *Psychooncology* 2011 Jun;20(6):655-663. [doi: [10.1002/pon.1939](https://doi.org/10.1002/pon.1939)] [Medline: [21381148](https://pubmed.ncbi.nlm.nih.gov/21381148/)]
37. Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM. How user characteristics affect use patterns in web-based illness management support for patients with breast and prostate cancer. *J Med Internet Res* 2013;15(3):e34 [FREE Full text] [doi: [10.2196/jmir.2285](https://doi.org/10.2196/jmir.2285)] [Medline: [23454601](https://pubmed.ncbi.nlm.nih.gov/23454601/)]
38. Han JY, Hawkins RP, Shaw BR, Pingree S, McTavish F, Gustafson DH. Unraveling Uses and Effects of an Interactive Health Communication System. *J Broadcast Electron Media* 2009;53(1):112-133 [FREE Full text] [doi: [10.1080/08838150802643787](https://doi.org/10.1080/08838150802643787)] [Medline: [23172985](https://pubmed.ncbi.nlm.nih.gov/23172985/)]

Abbreviations

BREATH: BREAst cancer eHealth

CBT: cognitive behavior therapy

HADS: Hospital Anxiety and Depression Scale

HADS-T: Hospital Anxiety and Depression Scale total

ICT: information and communication technology

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 05.02.13; peer-reviewed by S Kelders, L Donkin, R Kok; comments to author 07.03.13; revised version received 19.04.13; accepted 09.07.13; published 19.08.13.

Please cite as:

van den Berg SW, Peters EJ, Kraaijeveld JF, Gielissen MFM, Prins JB

Usage of a Generic Web-Based Self-Management Intervention for Breast Cancer Survivors: Substudy Analysis of the BREATH Trial
J Med Internet Res 2013;15(8):e170

URL: <http://www.jmir.org/2013/8/e170/>

doi: [10.2196/jmir.2566](https://doi.org/10.2196/jmir.2566)

PMID: [23958584](https://pubmed.ncbi.nlm.nih.gov/23958584/)

©Sanne W van den Berg, Esmee J Peters, J Frank Kraaijeveld, Marieke FM Gielissen, Judith B Prins. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 19.08.2013. 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

Pedometer-Based Internet-Mediated Intervention For Adults With Chronic Low Back Pain: Randomized Controlled Trial

Sarah L Krein^{1,2}, PhD, RN; Reema Kadri¹, MLIS; Maria Hughes¹, MPT; Eve A Kerr^{1,2}, MD, MPH; John D Piette^{1,2}, MSc, PhD; Rob Holleman¹, MPH; Hyungjin Myra Kim^{1,3}, ScD; Caroline R Richardson^{1,4}, MD

¹VA Ann Arbor Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, MI, United States

²Division of General Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States

³Center for Statistical Consultation and Research, University of Michigan, Ann Arbor, MI, United States

⁴Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:

Sarah L Krein, PhD, RN

VA Ann Arbor Center for Clinical Management Research

VA Ann Arbor Healthcare System

HSR&D (152)

PO Box 130170

Ann Arbor, MI, 48113

United States

Phone: 1 734 845 3621

Fax: 1 734 222 7503

Email: skrein@umich.edu

Abstract

Background: Chronic pain, especially back pain, is a prevalent condition that is associated with disability, poor health status, anxiety and depression, decreased quality of life, and increased health services use and costs. Current evidence suggests that exercise is an effective strategy for managing chronic pain. However, there are few clinical programs that use generally available tools and a relatively low-cost approach to help patients with chronic back pain initiate and maintain an exercise program.

Objective: The objective of the study was to determine whether a pedometer-based, Internet-mediated intervention can reduce chronic back pain-related disability.

Methods: A parallel group randomized controlled trial was conducted with 1:1 allocation to the intervention or usual care group. 229 veterans with nonspecific chronic back pain were recruited from one Department of Veterans Affairs (VA) health care system. Participants randomized to the intervention received an uploading pedometer and had access to a website that provided automated walking goals, feedback, motivational messages, and social support through an e-community (n=111). Usual care participants (n=118) also received the uploading pedometer but did not receive the automated feedback or have access to the website. The primary outcome was measured using the Roland Morris Disability Questionnaire (RDQ) at 6 months (secondary) and 12 months (primary) with a difference in mean scores of at least 2 considered clinically meaningful. Both a complete case and all case analysis, using linear mixed effects models, were conducted to assess differences between study groups at both time points.

Results: Baseline mean RDQ scores were greater than 9 in both groups. Primary outcome data were provided by approximately 90% of intervention and usual care participants at both 6 and 12 months. At 6 months, average RDQ scores were 7.2 for intervention participants compared to 9.2 for usual care, an adjusted difference of 1.6 (95% CI 0.3-2.8, $P=.02$) for the complete case analysis and 1.2 (95% CI -0.09 to 2.5, $P=.07$) for the all case analysis. A post hoc analysis of patients with baseline RDQ scores ≥ 4 revealed even larger adjusted differences between groups at 6 months but at 12 months the differences were no longer statistically significant.

Conclusions: Intervention participants, compared with those receiving usual care, reported a greater decrease in back pain-related disability in the 6 months following study enrollment. Between-group differences were especially prominent for patients reporting greater baseline levels of disability but did not persist over 12 months. Primarily, automated interventions may be an efficient way to assist patients with managing chronic back pain; additional support may be needed to ensure continuing improvements.

Trial Registration: ClinicalTrials.gov NCT00694018; <http://clinicaltrials.gov/ct2/show/NCT00694018> (Archived by WebCite at <http://www.webcitation.org/6IsG4Y90E>).

KEYWORDS

chronic pain; Internet; randomized controlled trial; exercise therapy

Introduction

Low back pain is a significant health problem with approximately one-half of adults reporting back pain during a given year [1-3]. Low back pain that persists for longer than 3 months is considered chronic [4], and the longer the pain persists the greater the risk for long-term disability [5]. Chronic back pain is associated with functional limitations, social isolation, unemployment, and lost productivity [5-7], making it one of the most costly health conditions in the United States [8-11].

Exercise therapy has proven benefits for managing chronic back pain [12-14]. Specifically, exercise can prevent recurrence, reduce pain, improve function, and decrease disability for patients with chronic back pain [12,13,15-19]. It is also generally recognized that, to be effective, patients have to be willing and able to perform the recommended exercise and for continuing benefits remain adherent to the exercise program [18,20,21]. However, there are few efficient and effective strategies to help patients engage in exercise therapy for managing their chronic low back pain.

Internet-based programs are an increasingly popular option for promoting healthy behaviors, such as those related to diet and exercise, and for delivering behavior change interventions [22-24]. Studies have shown that the Internet can be used to successfully promote weight loss [25], increase physical activity [26], and improve patient self-activation [27] or self-management behaviors [22]. Studies of Internet-based interventions for pain, while somewhat limited, show a generally positive effect on pain levels and, to some extent, activity [27-30]. No studies, however, have focused primarily on exercise to reduce pain-related disability and improve patient function.

We conducted a randomized trial to investigate whether a pedometer-based, Internet-mediated intervention designed to assist patients with initiating and maintaining a regular walking program would reduce pain-related disability and functional interference among patients with chronic back pain at 6 months and over a 12-month timeframe.

Methods

Design Overview

We conducted a parallel group randomized controlled trial with participants allocated in a 1:1 ratio to the intervention or enhanced usual care (NCT00694018). This research was approved by the Department of Veterans Affairs (VA), Ann Arbor Healthcare System institutional review board. The study protocol, including conceptual framework, is described in detail elsewhere [31], with key elements summarized below. There were no significant changes in methods following study initiation.

Setting and Participants

Participants were recruited from one VA Healthcare System between May 2009 and March 2011. Eligible participants were aged 18 years or older and identified through provider referrals to back class and use of the VA electronic medical record system. Specifically, we identified patients who had 2 or more outpatient encounters in the previous 12 months with a diagnosis of back pain with no neurologic findings (ICD-9-CM codes 724.2, 724.5, 846.0-846.9).

Study staff used a standardized protocol to screen potential participants by phone or, for a minority of patients who could not be reached by phone, in person when they arrived for back class. Eligibility criteria included: (1) persistent back pain >3 months, (2) self-reported sedentary lifestyle (defined as <150 minutes of physical activity per week in accordance with the US Department of Health and Human Services 2008 Physical Activity Guidelines for Americans [32]), (3) weekly access to a computer with a USB port and Internet access, (4) ability to provide written informed consent and communicate in English, (5) community residence, (6) ability to walk at least one block, and (7) report they are not pregnant. Prior to participation, all eligible patients had to attend back class and obtain medical clearance. Back class, led by a physical therapist, provided general education about managing back pain. Participants also performed back-specific strengthening and stretching exercises under the supervision of a physical therapist.

Eligible participants then attended a study enrollment session at which time they provided written informed consent and were told they were helping test an Internet-based program and would be assigned to one of two groups: (1) an enhanced care group that would upload pedometer data weekly and have access to a study website and computer discussion group (Internet support group), or (2) a usual care group that would upload pedometer data monthly (monthly upload group). All participants received an uploading pedometer (the Omron HJ-720ITC, which stores 42 days of step-count data and has an embedded USB port [33]), along with general guidance on using the pedometer and instructions for logging onto and uploading data to the study website. To establish a baseline step count that was not influenced by use of the pedometer information, participants were instructed to wear their pedometer for 7 days with the display covered before completing their first upload.

Randomization

After completing the baseline survey, uploading 7 days of useable pedometer data, and receiving medical clearance, each participant was randomly allocated in a 1:1 ratio to the intervention or usual care group by a computer program (using a random number generator). The program also generated an email message to inform participants about their group assignment (Internet support or monthly upload) and instructions to remove the sticker covering the pedometer display.

Intervention

The study intervention, based on the Stepping Up to Health program [31,34], consisted of three primary components: (1) the uploading pedometer, (2) a website that provided automated goal setting and feedback, targeted messages, and educational materials, and (3) an e-community [31] (see [Multimedia Appendix 1](#)). The conceptual framework and more detailed description of the intervention components are published elsewhere [31]. Briefly, participants were instructed to wear their pedometer from the time they got up in the morning until they went to bed. Intervention participants then received weekly email reminders to upload their pedometer data, which was used to establish weekly individualized walking goals. Each participant's goal was based on their average total step count in the prior week with a fixed number of steps (800) added to promote a gradual increase in walking for the following week. The step count goal was emailed to the participant each week and posted on the study website.

The study website, which was fully accessible to intervention participants, also included graphical and written feedback about their progress toward their walking goals and contained pain- or activity-related motivational and informational messages. These messages included quick tips, which changed every other day, and weekly updates about topics in the news. Back class materials, which included handouts about topics such as body mechanics, use of cold packs, lumbar rolls, and good posture, as well as a video demonstrating specific strengthening and stretching exercises were also available on the website. Finally, the website based e-community or forum allowed participants to post suggestions, ask questions, and share stories. Topics discussed included mental health concerns, such as depression, strategies for walking such as walking the dog or interesting hiking trails, walking during hot weather and cold weather, and use of alternative pain management strategies such as massage. Research staff participated in and monitored the forum posts as well as used the forum as a venue to generate competitions to encourage meeting walking goals.

Enhanced Usual Care

Usual care participants also received the uploading pedometer and monthly email reminders to upload their pedometer data. However, they did not receive any goals or feedback and their access to the study website was limited to completing surveys and reporting adverse events only.

Monitoring of Adverse Events

Both groups were encouraged to report any health problems via the website, email, or phone. Four weeks after randomization and every 8 weeks thereafter, participants were prompted to complete a survey that asked about specific adverse events (eg, heart attack) and symptoms such as shortness of breath. This information was closely monitored and participants with potentially serious health-related problems were contacted for further assessment and follow-up.

Outcomes and Follow-Up

Outcomes were measured at baseline, 6 months, and 12 months using a survey administered through the study website, or by a mailed questionnaire if the participant could not complete the

computerized instrument. The prespecified primary outcome was pain-related disability at 12 months, as measured using the back pain-specific Roland Morris Disability Questionnaire (RDQ) [35], and a generic pain-related function measure from the Medical Outcomes Study (MOS) [36]. The RDQ, a 24-item scale with higher scores indicating greater disability, has been widely used in back pain studies as a measure of self-perceived disability [35,37-39]. The MOS measure assesses the effect of pain on mood and behaviors as well as pain severity, with higher scores also indicating greater functional interference [36].

Pain intensity, a secondary outcome, was evaluated using a numeric rating scale with standard anchors (0="no pain" and 10="worst pain imaginable") [40]. Walking, also a secondary outcome, was measured as the average number of steps per day over the past 7 days using step-count data collected through the pedometer uploads. Other secondary outcomes included pain-related fear-avoidance, measured using the Fear-Avoidance Beliefs Questionnaire physical activity subscale (higher scores reflect higher levels of fear-avoidance) [41], and self-efficacy for exercise, measured using the Exercise Regularly Scale, with higher scores indicating higher levels of self-efficacy [42]. Additional data collected at baseline included age, gender, race, employment status, education level, relationship status, average household income, body mass index, and use of narcotic medications for pain management. An administrative interface to the website provided data on the number of pedometer uploads and website log-ins.

Sample Size

Sample size was based on the RDQ score as the primary outcome with a minimally detectable and clinically meaningful effect size determined as a difference of 0.4 standard deviation (SD) in change scores or a 2-point difference, based on published data [38,43,44]. To detect a difference of 0.4 SD with 80% power using a two-sided 0.05 level 2 group *t* test, we sought to enroll 130 subjects in each group, to allow for an attrition rate of 25% at 1 year.

Statistical Analysis

The analyst assessing final trial outcomes was blinded to study assignment. All analyses were conducted using an intent-to-treat approach with participants analyzed according to original group assignment. We conducted both complete and all case analyses to assess differences between groups in change in RDQ at 6 and 12 months. The complete case analysis was conducted using multiple linear regression models with adjustment for baseline values of the RDQ. The all case analysis was conducted using linear mixed-effects models, allowing us to use data from all participants and provide an unbiased estimate of the outcome, assuming data are missing at random [45]. For example, for our 12-month analysis, RDQ scores at baseline and 12 months were used as dependent variables, with the primary independent variables consisting of an indicator for the intervention group and an interaction term of time by intervention group. Each participant's data was modeled using a random intercept to allow within-patient correlation of the repeated measures. Adjustment for covariates was only planned if an imbalance was found between groups at baseline.

We also conducted a post hoc subgroup analysis of participants with baseline RDQ scores of ≥ 4 . As a pragmatic trial we did not screen based on RDQ scores, and some participants had baseline scores that were very low or even 0. Thus, to assess the effect of the intervention on participants reporting at least modest levels of back pain-related disability at baseline, we conducted a subgroup analysis of those with baseline RDQ scores of ≥ 4 using the same methods previously described.

Analyses were conducted using Stata 11.2 and all reported *P* values are from adjusted analyses.

Results

Summary

Over 1400 potential participants (Figure 1) were assessed for eligibility. Primary reasons for ineligibility were lack of regular access to a computer or the Internet ($n=310$) and being too physically active ($n=159$). Of those determined to be eligible, 229 completed all of the steps in the enrollment process, with 111 randomly allocated to the Internet-mediated intervention and 118 to enhanced usual care. Primary outcome data were provided by 91% of intervention and 90% of usual care participants at 6 months, and by 92% of those in the intervention group and 89% receiving usual care at 12 months.

Baseline Characteristics

Participants were predominantly male and white, with an average age of 51 years (Table 1). The majority had completed some college, were either married or living with someone as a couple, and the mean body mass index was over 30. At baseline, less than 40% of participants reported being employed full- or part-time and over 40% reported taking narcotic medications for their back pain. None of the observed differences in baseline characteristics were statistically significant.

Primary Outcomes

At baseline, mean RDQ scores were greater than 9 in both groups (Table 1), indicating moderately severe back pain-related disability. The mean RDQ score at 6 months was 7.2 for intervention participants compared to 9.2 for those in usual care (Figure 2), an adjusted difference of 1.6 (95% CI 0.3-2.8, $P=.02$) for the complete case analysis and 1.2 (95% CI -0.09 to 2.5, $P=.07$) for the all case analysis (Table 2). When restricted to the subgroup with at least moderate back pain at baseline (RDQ score ≥ 4) (Figure 2, Table 2), patients in the intervention had a significant improvement in back pain-related disability compared to the control group, an adjusted difference of approximately 2 in both the complete (1.9, 95% CI 0.5-3.3, $P=.01$) and all case (1.7, 95% CI 0.3-3.0, $P=.02$) analyses. RDQ scores continued to decline between 6 and 12 months in both groups and, while scores for the intervention group remained lower than for usual care, at 12 months these differences were

no longer statistically significant. The MOS function measure also suggested greater improvements in function for intervention compared to usual care participants at 6 months (Figure 2), but none of the adjusted differences were statistically significantly different.

Secondary Outcomes

At baseline, pain severity was rated at approximately 6 on a 0-10 scale by both intervention and usual care participants (Table 1). Reported pain levels decreased in both groups at 6 months and remained lower than baseline at 12 months. The greatest change occurred between baseline and 6 months among those in the intervention group (6.0-4.7 vs 6.1-5.2 in the control group), although the adjusted difference between arms of 0.5 was not significant (Table 3).

Average step counts of slightly more than 4000 steps per day at baseline in each group increased at 6 months for intervention patients, with an adjusted difference between groups of more than 700 steps. By 12 months, however, the adjusted difference between groups was only 100-200 steps. Exercise self-efficacy scores appeared to be the same or lower (worse) for both groups at 6 months, although the decrease was significantly less for those in the intervention compared to the control group, an adjusted difference of 0.8 (95% CI 0.24-1.4, $P=.01$) in the complete case analysis and 0.7 (95% CI 0.12-1.2, $P=.02$) for the all case analysis (Table 3). This difference did not persist at 12 months. There was no difference between groups in the physical activity fear-avoidance scale at any time point.

Intervention Engagement

Intervention participants uploaded pedometer data at least once per week for a median of 32 weeks (62% of the recommended time), although more than 25% of participants uploaded data for at least 42 weeks (80% compliance). However, intervention participants logged into the website at least once per week for a median of only 20 weeks (38% of the recommended time), with approximately 20% logging in for at least 42 weeks.

Adverse Events

During the study, approximately 600 adverse events were reported by participants (250 by those in usual care and nearly 350 by those in the intervention). These events ranged from calluses to chest pain. Worsening back pain, the most frequently reported event, accounted for 29% of events reported by the usual care group and 25% of those reported by the intervention group. Overall, more musculoskeletal events ($n=112$) were reported than cardiovascular events ($n=85$), and musculoskeletal injuries were more likely to be reported by participants in the intervention group compared to those in usual care. However, no major study-related adverse events (eg, heart attack) were identified for either group.

Table 1. Participant baseline characteristics.

Characteristic	Internet-mediated inter- vention (n=111)	Enhanced usual care (n=118)
Age (y), mean (SD)	51.2 (12.5)	51.9 (12.8)
Male (%)	89	86
Race (%)		
White	74	86
Black	13	9
Other or prefer not to answer	14	5
Education level (%)		
High school or less	29	25
Some college	56	59
4 years of college or more	16	16
Married or living with a partner (%)	59	68
Employed full-time or part-time (%)	39	31
Annual household income (%)		
<US \$10,000	18	13
US \$10,000-\$39,999	61	54
≥ US \$40,000	21	33
Take narcotic medications for back pain (%)	41	49
General health status, fair or poor (%)	41	43
Body mass index, mean (SD)	30.6 (5.7)	31.6 (5.5)
RDQ score (0-24) ^a , mean (SD)	9.1 (6.0)	9.8 (5.7)
MOS pain-related functional interference score (0-100) ^a , mean (SD)	48.5 (18.6)	51.8 (16.3)
Level of pain severity, 0-10 scale ^a , mean (SD)	6.0 (1.9)	6.1 (1.6)
Daily step counts, mean (SD)	4492.9 (2749.9)	4321.9 (2285.4)
Exercise self-efficacy score, 1-10 ^b , mean (SD)	6.8 (2.1)	6.5 (2.3)
Physical activity fear-avoidance behavior scale, 0-28 ^a , mean (SD)	13.9 (5.9)	15.1 (6.0)

^alower scores are better^bhigher scores are better

Figure 1. Study flow diagram.

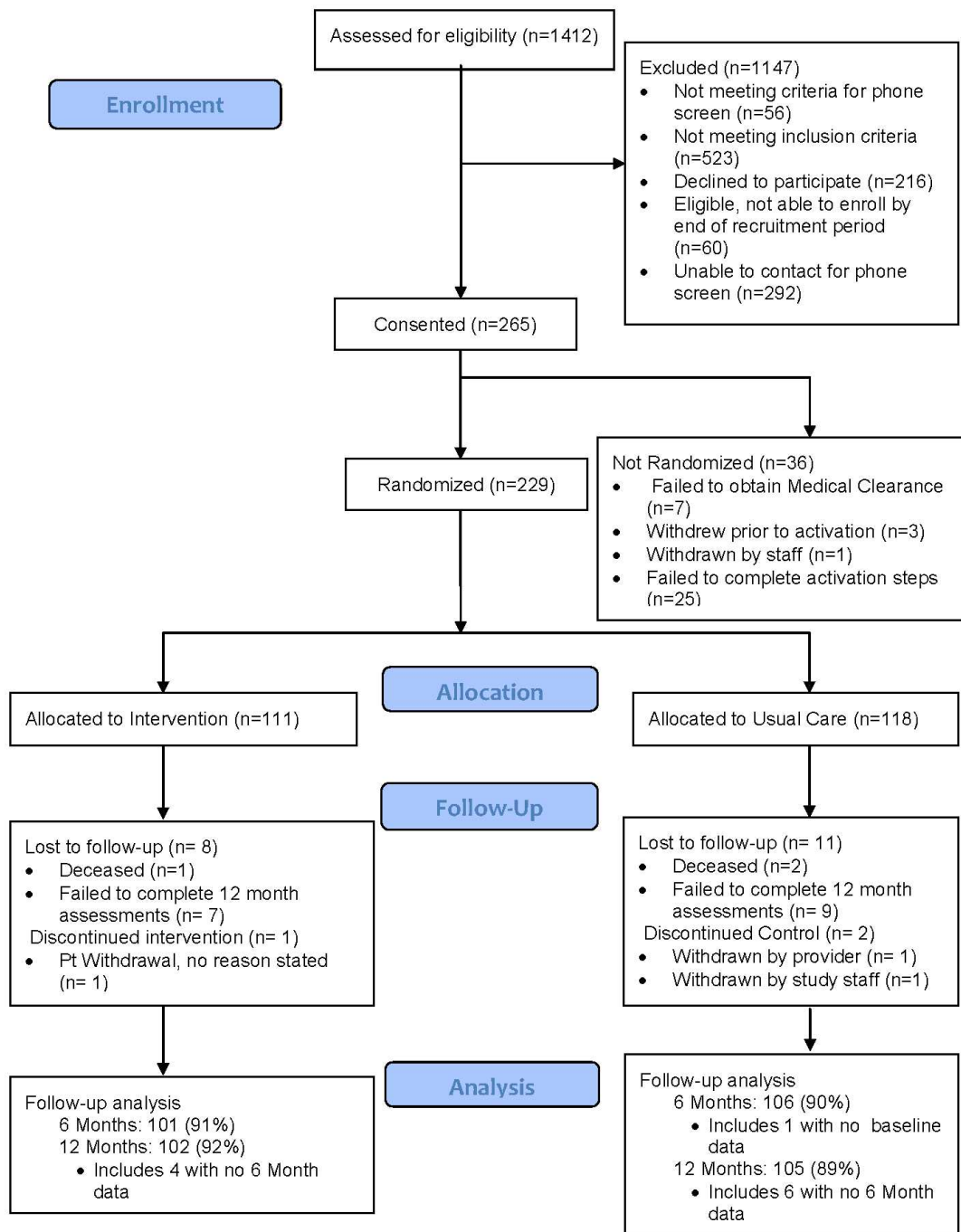


Table 2. Primary outcomes, back pain-specific and general pain-related function.

Primary outcome	Adjusted between-group difference ^a (95% CI)			
	Complete case	<i>P</i> value	All case	<i>P</i> value
RDQ score (0-24)				
6 months	1.6 (0.3 to 2.8)	.02	1.2 (-0.09 to 2.5)	.07
12 months	1.2 (-0.3 to 2.7)	.11	0.7 (-0.8 to 2.2)	.38
MOS pain-related functional interference score (0-100)				
6 months	3.6 (-0.51 to 7.7)	.09	2.5 (-1.5 to 6.5)	.23
12 months	0.1 (-4.0 to 4.2)	.97	-1.4 (-5.4 to 2.5)	.48
Subgroup with RDQ scores ≥ 4 at baseline RDQ score (0-24)				
6 months	1.9 (0.5 to 3.3)	.01	1.7 (0.3 to 3.0)	.02
12 months	1.1 (-0.6 to 2.7)	.20	0.8 (-0.8 to 2.4)	.34
MOS pain-related functional interference score (0-100)				
6 months	4.6 (-0.1 to 9.3)	.05	3.8 (-0.7 to 8.3)	.10
12 months	-0.5 (-5.0 to 4.0)	.83	-1.5 (-5.8 to 2.8)	.49

^aAdjusted for baseline values and calculated as pain or function in enhanced usual care group minus Internet-mediated intervention group so that positive scores reflect greater improvement in the intervention group.

Table 3. Secondary outcomes.

Secondary outcome	Mean (SD)	Adjusted between-group difference ^a (95% CI)				
		Internet-mediated intervention	Enhanced usual care	Complete case	<i>P</i> value	All case
Level of pain severity, 0-10 scale^b						
6 months	4.7 (2.1)	5.2 (2.1)	0.5 (-0.01 to 0.98)	.06	0.5 (-0.03 to 0.9)	.07
12 months	5.4 (2.2)	5.6 (2.0)	0.1 (-0.4 to 0.5)	.81	0.04 (-0.4 to 0.5)	.86
Daily step counts^{c,d}						
6 months	5370.0 (3180.8)	4682.5 (2925.0)	725.5 (-193.6 to 1644.7)	.12	724.0 (-75.2 to 1523.2)	.08
12 months	4681.8 (3000.6)	4758.1 (2991.1)	122.4 (-623.9 to 868.6)	.75	143.4 (-460.2 to 747.1)	.64
Exercise self-efficacy score^c						
6 months	6.7 (2.4)	5.7 (2.5)	0.8 (0.24 to 1.4)	.01	0.7 (0.12 to 1.2)	.02
12 months	6.4 (2.6)	5.9 (2.3)	0.3 (-0.3 to 0.9)	.32	0.2 (-0.4 to 0.74)	.55
Physical activity fear-avoidance behavior scale^b						
6 months	13.2 (6.0)	14.0 (5.9)	0.6 (-0.88 to 2.1)	.42	-0.1 (-1.6 to 1.5)	.94
12 months	13.3 (6.7)	15.1 (6.1)	1.1 (-0.5 to 2.7)	.18	0.6 (-1.1 to 2.2)	.50

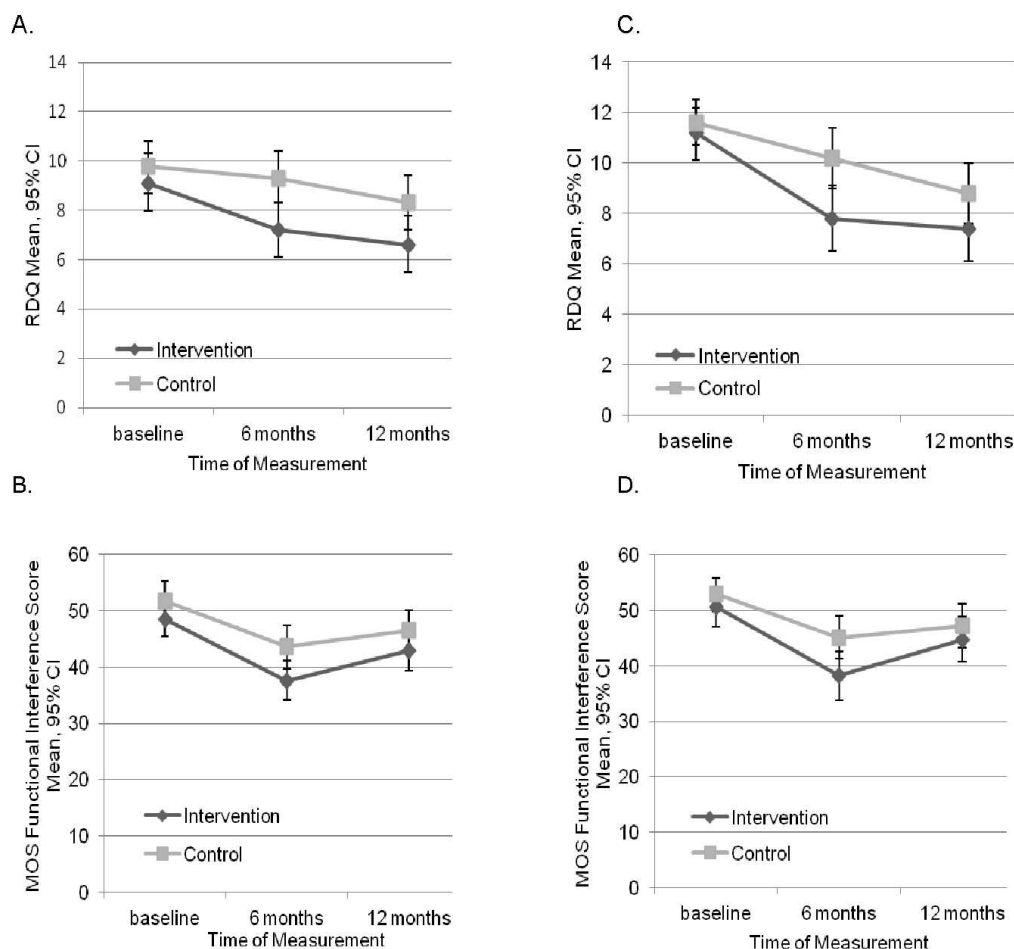
^aAdjusted for baseline values and calculated as pain or function in enhanced usual care group minus Internet-mediated intervention group so that positive scores reflect greater improvement in the intervention group.

^blower scores are better

^chigher scores are better

^dPedometer data: intervention (n=84 at 6 months, n=78 at 12 months), usual care (n=70 at 6 months, n=68 at 12 months).

Figure 2. Mean RDQ scores (top) and MOS functional interference scores (bottom). A and B: full sample, C and D: patients with baseline RDQ scores ≥ 4 .



Discussion

Principal Findings

Improving management of chronic pain is a significant public health challenge and moral imperative according to a recent Institute of Medicine report [8]. More than 1 million adults in the United States have chronic pain, with low back pain being the most frequently reported condition [8]. Our findings show that an automated, Internet-mediated walking intervention may help to reduce back pain-related disability among patients with chronic back pain, although the benefits did not persist for the entire 12-month study period. Improvement was greatest for those individuals reporting moderate to severe levels of pain-related disability at baseline.

The functional results observed are generally similar to those found in other recent studies of non-invasive interventions, such as yoga and massage [46,47]. These studies also tend to show more rapid improvements for those receiving the intervention but with gradual improvements over time for those in usual care. Moreover, if we employ the criteria proposed by Jordan

and colleagues [48] to classify patients as clinically improved or at least possibly improved (compared with not improved), as defined by a reduction in the RDQ score of at least 30% at 6 months, we find that 46% of those in the intervention versus 27% in the control group would meet this definition. Although we did not have a global health question and so are unable to isolate what proportion would qualify as definitely improved, this classification generally corresponds with other measures that suggest clinical improvement, such as return to work, less pain, improved function, and fewer physician visits [48]. Thus, we believe that our findings suggest that automated, remotely delivered interventions can be effectively used to promote a more rapid reduction in back pain-related disability and supplement care for patients with chronic low back pain. Further investigation is needed, however, to understand the characteristics of patients who had an early or enduring response to the intervention so that we may better target patients most likely to benefit and broaden the response.

Given the proven benefits of exercise for managing low back pain [19], a key component of the intervention focused on increasing daily step counts (ie, walking). During the first 6

months of the study, we saw an increase of nearly 700 more steps or one-third of a mile per day among intervention compared to usual care participants. Although not a statistically significant difference, we believe that even modest increases in activity can be beneficial. As one intervention participant noted: "I didn't know what the walking could do for me. But...it seemed to alleviate my back pain...the true test came when I had to go off the program because of my illness and the back pain returned. In fact, just up until recently when I had resumed walking." On the other hand, step-count improvements were not sustained for the entire 12 months and poor adherence or declining engagement, as demonstrated by the percentage of patients who uploaded or logged into the website, could in part account for the lack of sustained benefit or added improvement over time. Although we do not know specific reasons for this lack of participation, these data suggest that additional strategies to keep people active and engaged may be needed. This could include, for example, an online coaching component, which has been shown to improve adherence to other types of behavioral changes [47-49].

Our monitoring of adverse events showed a higher number of reported events by intervention participants. This information was, however, collected solely through self-report and we expect that some of the difference in the overall number of events reported between groups could be due to our more frequent contact with intervention participants via email and through the website. In addition, despite the higher level of musculoskeletal events reported by intervention participants, we found no evidence that the intervention led to excessive harms. Thus, even though more work to understand the circumstances for those reporting musculoskeletal problems or worsening back pain may be required, these findings add to the evidence base to support walking as a generally safe and potentially effective intervention for some patients with chronic low back pain [49-52].

Other potential mechanisms of action are less clear. Despite a marginally greater decrease in pain levels among intervention participants at 6 months, this effect did not persist at 12 months. In addition, while there was a significant difference between groups in self-efficacy for exercise at 6 months, rather than the hypothesized improvement for those in the intervention, both groups reported lower levels of self-efficacy. However, the decline was smaller for those receiving the intervention. The

reason for the decrease is not entirely clear but may be largely due to an unrealistic assessment of self-efficacy at baseline [53].

Limitations

Among the strengths of our study are the high rate of participant follow-up and our collection of detailed adverse event information. This study also has several limitations. First, patients were recruited from only 1 medical center and the sample was predominantly male. Although more than 10% of participants were female, which is relatively high for studies using a general VA patient population, the number is not sufficient for a formal subgroup analysis. However, based on trials of similar types of interventions, we expect this approach could be even more effective among women [54]. Second, we are not able to directly compare our results to other types of back pain interventions (eg, yoga), although as previously noted the general trajectory of our primary outcome (RDQ score) appears consistent with recent trials in this area. Third, although a consistent data collection format is generally recommended [55], we used both Internet-based and paper surveys. However, prior research has demonstrated similar psychometric properties between Internet and paper-and-pencil questionnaires [55] and specifically equivalence for our primary outcome [56]. We also believe that using both modes helped to ensure a high follow-up rate. Finally, as a multifaceted intervention, we are not able to determine which elements were most effective and can only draw conclusions about the program as a whole. Nonetheless, our results highlight the importance of providing active support (eg, goal setting and feedback) to encourage walking as compared with simply giving someone a pedometer to track step counts.

Conclusions

In sum, our findings indicate that a facilitated walking intervention that uses an uploading pedometer and the Internet may help to reduce back pain-related disability among patients with chronic back pain, at least in the short term. Additional support, however, is likely needed to ensure continuing improvements long term. Nevertheless, this type of primarily automated intervention can be used to deliver care with broad reach and could be an efficient way of delivering or supplementing care provided through traditional facility-based programs.

Acknowledgments

This project was funded through a grant from the Department of Veterans Affairs, Health Services Research and Development Service (IIR 07-177). The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government. The study sponsors had no role in the design or conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript. Dr Piette is a VA Senior Research Career Scientist. Dr Richardson was supported in part by a career development award from the National Heart, Lung, and Blood Institute (K23HL075098) and a Physician Faculty Scholars Program award from the Robert Wood Johnson Foundation (57408).

Development of the Stepping Up to Health intervention platform was supported by pilot grant funding from the following University of Michigan centers: Michigan Diabetes Research and Training Center (P60 DK020572), the Center for Health Communications Research (P50 CA101451), and the Michigan Institute for Clinical and Health Research (NIH #UL1RR024986). Technical development of the Stepping Up to Health website could not have been accomplished without the expertise of our Web

developers, Michael Hess, with assistance from Elizabeth Wilson and Adrienne Janney. Thanks also to Jill Bowdler for assistance with manuscript preparation. Finally, the authors would like to thank Robert Kerns, Jack Rosenberg, and Robert Werner for their assistance during the design of this study, as well as the Physical Therapy Service at the VA Ann Arbor Medical Center for their help during the conduct of the project.

Authors' Contributions

SLK, JDP, EAK, HMK, and CRR were involved in the conception and design of the study. RK and MH were involved in study data collection. SLK, RH, and HMK conducted study analyses. All authors were involved in drafting the manuscript and revising it for critically important content. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Intervention components.

[[PDF File \(Adobe PDF File\), 915KB - jmir_v15i8e181_app1.pdf](#)]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [57].

[[PDF File \(Adobe PDF File\), 1000KB - jmir_v15i8e181_app2.pdf](#)]

References

1. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine (Phila Pa 1976)* 2006 Nov 1;31(23):2724-2727. [doi: [10.1097/01.brs.0000244618.06877.cd](https://doi.org/10.1097/01.brs.0000244618.06877.cd)] [Medline: [17077742](#)]
2. Brooks PM. The burden of musculoskeletal disease--a global perspective. *Clin Rheumatol* 2006 Nov;25(6):778-781. [doi: [10.1007/s10067-006-0240-3](https://doi.org/10.1007/s10067-006-0240-3)] [Medline: [16609823](#)]
3. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini EH, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998 May;41(5):778-799. [doi: [10.1002/1529-0131\(199805\)41:5<778::AID-ART4>3.0.CO;2-V](https://doi.org/10.1002/1529-0131(199805)41:5<778::AID-ART4>3.0.CO;2-V)] [Medline: [9588729](#)]
4. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006 Jun 17;332(7555):1430-1434 [FREE Full text] [doi: [10.1136/bmj.332.7555.1430](https://doi.org/10.1136/bmj.332.7555.1430)] [Medline: [16777886](#)]
5. Andersson GB. Epidemiological features of chronic low-back pain. *Lancet* 1999 Aug 14;354(9178):581-585. [doi: [10.1016/S0140-6736\(99\)01312-4](https://doi.org/10.1016/S0140-6736(99)01312-4)] [Medline: [10470716](#)]
6. Weiner DK, Haggerty CL, Kritchevsky SB, Harris T, Simonsick EM, Nevitt M, Health, Aging, and Body Composition Research Group. How does low back pain impact physical function in independent, well-functioning older adults? Evidence from the Health ABC Cohort and implications for the future. *Pain Med* 2003 Dec;4(4):311-320. [Medline: [14750907](#)]
7. Jacobs JM, Hammerman-Rozenberg R, Cohen A, Stessman J. Chronic back pain among the elderly: prevalence, associations, and predictors. *Spine (Phila Pa 1976)* 2006 Apr 1;31(7):E203-E207. [doi: [10.1097/01.brs.0000206367.57918.3c](https://doi.org/10.1097/01.brs.0000206367.57918.3c)] [Medline: [16582841](#)]
8. Institute of Medicine (US) Committee on Advancing Pain Research Care and Education. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington, DC: National Academies Press; 2011.
9. Soni A. Back problems: Use and expenditures for the U.S. adult population, 2007. 2010. URL: http://meps.ahrq.gov/mepsweb/data_files/publications/st289/stat289.pdf [accessed 2013-02-27] [WebCite Cache ID 6EkNbJzhX]
10. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, et al. The rising prevalence of chronic low back pain. *Arch Intern Med* 2009 Feb 9;169(3):251-258. [doi: [10.1001/archinternmed.2008.543](https://doi.org/10.1001/archinternmed.2008.543)] [Medline: [19204216](#)]
11. Druss BG, Marcus SC, Olfson M, Pincus HA. The most expensive medical conditions in America. *Health Aff (Millwood)* 2002 Aug;21(4):105-111 [FREE Full text] [Medline: [12117121](#)]
12. Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. *Ann Intern Med* 2005 May 3;142(9):765-775. [Medline: [15867409](#)]
13. van Tulder MW, Koes B, Malmivaara A. Outcome of non-invasive treatment modalities on back pain: an evidence-based review. *Eur Spine J* 2006 Jan;15 Suppl 1:S64-S81 [FREE Full text] [doi: [10.1007/s00586-005-1048-6](https://doi.org/10.1007/s00586-005-1048-6)] [Medline: [16320031](#)]
14. Chou R, Huffman LH, American Pain Society, American College of Physicians. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. *Ann Intern Med* 2007 Oct 2;147(7):492-504. [Medline: [17909210](#)]
15. Sherman KJ, Cherkin DC, Erro J, Miglioretti DL, Deyo RA. Comparing yoga, exercise, and a self-care book for chronic low back pain: a randomized, controlled trial. *Ann Intern Med* 2005 Dec 20;143(12):849-856. [Medline: [16365466](#)]

16. Von Korff M, Balderson BH, Saunders K, Miglioretti DL, Lin EH, Berry S, et al. A trial of an activating intervention for chronic back pain in primary care and physical therapy settings. *Pain* 2005 Feb;113(3):323-330. [doi: [10.1016/j.pain.2004.11.007](https://doi.org/10.1016/j.pain.2004.11.007)] [Medline: [15661440](https://pubmed.ncbi.nlm.nih.gov/15661440/)]
17. Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, et al. Active rehabilitation for chronic low back pain: cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. *BMC Musculoskelet Disord* 2006;7:5 [FREE Full text] [doi: [10.1186/1471-2474-7-5](https://doi.org/10.1186/1471-2474-7-5)] [Medline: [16426449](https://pubmed.ncbi.nlm.nih.gov/16426449/)]
18. Mailloux J, Finno M, Rainville J. Long-term exercise adherence in the elderly with chronic low back pain. *Am J Phys Med Rehabil* 2006 Feb;85(2):120-126. [Medline: [16428902](https://pubmed.ncbi.nlm.nih.gov/16428902/)]
19. van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *Eur Spine J* 2011 Jan;20(1):19-39 [FREE Full text] [doi: [10.1007/s00586-010-1518-3](https://doi.org/10.1007/s00586-010-1518-3)] [Medline: [20640863](https://pubmed.ncbi.nlm.nih.gov/20640863/)]
20. Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 2005 May 3;142(9):776-785. [Medline: [15867410](https://pubmed.ncbi.nlm.nih.gov/15867410/)]
21. Friedrich M, Gittler G, Arendasy M, Friedrich KM. Long-term effect of a combined exercise and motivational program on the level of disability of patients with chronic low back pain. *Spine (Phila Pa 1976)* 2005 May 1;30(9):995-1000. [Medline: [15864148](https://pubmed.ncbi.nlm.nih.gov/15864148/)]
22. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. *Med Care* 2006 Nov;44(11):964-971. [doi: [10.1097/01.mlr.0000233678.80203.c1](https://doi.org/10.1097/01.mlr.0000233678.80203.c1)] [Medline: [17063127](https://pubmed.ncbi.nlm.nih.gov/17063127/)]
23. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. *Arch Intern Med* 2006;166(15):1620-1625. [doi: [10.1001/archinte.166.15.1620](https://doi.org/10.1001/archinte.166.15.1620)] [Medline: [16908795](https://pubmed.ncbi.nlm.nih.gov/16908795/)]
24. 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/)]
25. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obes Rev* 2010 Apr;11(4):306-321. [doi: [10.1111/j.1467-789X.2009.00646.x](https://doi.org/10.1111/j.1467-789X.2009.00646.x)] [Medline: [19754633](https://pubmed.ncbi.nlm.nih.gov/19754633/)]
26. van den Berg MH, Schoones JW, Vliet Vlieland TP. Internet-based physical activity interventions: a systematic review of the literature. *J Med Internet Res* 2007;9(3):e26 [FREE Full text] [doi: [10.2196/jmir.9.3.e26](https://doi.org/10.2196/jmir.9.3.e26)] [Medline: [17942388](https://pubmed.ncbi.nlm.nih.gov/17942388/)]
27. 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/)]
28. Schulz PJ, Rubinell S, Hartung U. An internet-based approach to enhance self-management of chronic low back pain in the Italian-speaking population of Switzerland: results from a pilot study. *Int J Public Health* 2007;52(5):286-294. [Medline: [18030944](https://pubmed.ncbi.nlm.nih.gov/18030944/)]
29. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of Web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. *J Pain* 2010 Oct;11(10):917-929. [doi: [10.1016/j.jpain.2010.06.005](https://doi.org/10.1016/j.jpain.2010.06.005)] [Medline: [20650691](https://pubmed.ncbi.nlm.nih.gov/20650691/)]
30. Bender JL, Radhakrishnan A, Diorio C, Englesakis M, Jadad AR. Can pain be managed through the Internet? A systematic review of randomized controlled trials. *Pain* 2011 Aug;152(8):1740-1750. [doi: [10.1016/j.pain.2011.02.012](https://doi.org/10.1016/j.pain.2011.02.012)] [Medline: [21565446](https://pubmed.ncbi.nlm.nih.gov/21565446/)]
31. Krein SL, Metreger T, Kadri R, Hughes M, Kerr EA, Piette JD, et al. Veterans walk to beat back pain: study rationale, design and protocol of a randomized trial of a pedometer-based internet mediated intervention for patients with chronic low back pain. *BMC Musculoskelet Disord* 2010;11:205 [FREE Full text] [doi: [10.1186/1471-2474-11-205](https://doi.org/10.1186/1471-2474-11-205)] [Medline: [20836856](https://pubmed.ncbi.nlm.nih.gov/20836856/)]
32. U.S. Department of Health and Human Services. 2008 Physical Activity Guidelines for Americans Summary. 2008. URL: <http://www.health.gov/paguidelines/guidelines/summary.aspx> [accessed 2013-05-30] [WebCite Cache ID 6H0g4ccVE]
33. Zhu W, Lee M. Invariance of wearing location of Omron-BI pedometers: a validation study. *J Phys Act Health* 2010 Nov;7(6):706-717. [Medline: [21088300](https://pubmed.ncbi.nlm.nih.gov/21088300/)]
34. Richardson CR, Mehari KS, McIntyre LG, Janney AW, Fortlage LA, Sen A, et al. A randomized trial comparing structured and lifestyle goals in an internet-mediated walking program for people with type 2 diabetes. *Int J Behav Nutr Phys Act* 2007;4:59 [FREE Full text] [doi: [10.1186/1479-5868-4-59](https://doi.org/10.1186/1479-5868-4-59)] [Medline: [18021411](https://pubmed.ncbi.nlm.nih.gov/18021411/)]
35. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)* 1983 Mar;8(2):141-144. [Medline: [6222486](https://pubmed.ncbi.nlm.nih.gov/6222486/)]
36. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-483. [Medline: [1593914](https://pubmed.ncbi.nlm.nih.gov/1593914/)]
37. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)* 2000 Dec 15;25(24):3115-3124. [Medline: [11124727](https://pubmed.ncbi.nlm.nih.gov/11124727/)]

38. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 2. *Phys Ther* 1998 Nov;78(11):1197-1207 [[FREE Full text](#)] [Medline: [9806624](#)]
39. Hsieh CY, Phillips RB, Adams AH, Pope MH. Functional outcomes of low back pain: comparison of four treatment groups in a randomized controlled trial. *J Manipulative Physiol Ther* 1992 Jan;15(1):4-9. [Medline: [1531488](#)]
40. Mularski RA, White-Chu F, Overbay D, Miller L, Asch SM, Ganzini L. Measuring pain as the 5th vital sign does not improve quality of pain management. *J Gen Intern Med* 2006 Jun;21(6):607-612 [[FREE Full text](#)] [doi: [10.1111/j.1525-1497.2006.00415.x](#)] [Medline: [16808744](#)]
41. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993 Feb;52(2):157-168. [Medline: [8455963](#)]
42. Lorig K, Stewart A, Ritter P, Gonzalez V, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks, CA: Sage Publications; 1996.
43. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine (Phila Pa 1976)* 1995 Sep 1;20(17):1899-908; discussion 1909. [Medline: [8560339](#)]
44. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002 Jan;82(1):8-24 [[FREE Full text](#)] [Medline: [11784274](#)]
45. Schafer JL, Graham JW. Missing data: our view of the state of the art. *Psychol Methods* 2002 Jun;7(2):147-177. [Medline: [12090408](#)]
46. Sherman KJ, Cherkin DC, Wellman RD, Cook AJ, Hawkes RJ, Delaney K, et al. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. *Arch Intern Med* 2011 Dec 12;171(22):2019-2026 [[FREE Full text](#)] [doi: [10.1001/archinternmed.2011.524](#)] [Medline: [22025101](#)]
47. Cherkin DC, Sherman KJ, Kahn J, Wellman R, Cook AJ, Johnson E, et al. A comparison of the effects of 2 types of massage and usual care on chronic low back pain: a randomized, controlled trial. *Ann Intern Med* 2011 Jul 5;155(1):1-9 [[FREE Full text](#)] [doi: [10.7326/0003-4819-155-1-201107050-00002](#)] [Medline: [21727288](#)]
48. Jordan K, Dunn KM, Lewis M, Croft P. A minimal clinically important difference was derived for the Roland-Morris Disability Questionnaire for low back pain. *J Clin Epidemiol* 2006 Jan;59(1):45-52. [doi: [10.1016/j.jclinepi.2005.03.018](#)] [Medline: [16360560](#)]
49. Hendrick P, Te Wake AM, TikkiSETTY AS, Wulff L, Yap C, Milosavljevic S. The effectiveness of walking as an intervention for low back pain: a systematic review. *Eur Spine J* 2010 Oct;19(10):1613-1620 [[FREE Full text](#)] [doi: [10.1007/s00586-010-1412-z](#)] [Medline: [20414688](#)]
50. Hartvigsen J, Morsø L, Bendix T, Manniche C. Supervised and non-supervised Nordic walking in the treatment of chronic low back pain: a single blind randomized clinical trial. *BMC Musculoskelet Disord* 2010;11:30 [[FREE Full text](#)] [doi: [10.1186/1471-2474-11-30](#)] [Medline: [20146793](#)]
51. Murtezani A, Hundozi H, Orovcanec N, Sllamniku S, Osmani T. A comparison of high intensity aerobic exercise and passive modalities for the treatment of workers with chronic low back pain: a randomized, controlled trial. *Eur J Phys Rehabil Med* 2011 Sep;47(3):359-366 [[FREE Full text](#)] [Medline: [21602759](#)]
52. Shnayderman I, Katz-Leurer M. An aerobic walking programme versus muscle strengthening programme for chronic low back pain: a randomized controlled trial. *Clin Rehabil* 2013 Mar;27(3):207-214. [doi: [10.1177/0269215512453353](#)] [Medline: [22850802](#)]
53. Bandura A. *Self-efficacy: the exercise of control*. New York: W.H. Freeman; 1997.
54. Johnson F, Wardle J. The association between weight loss and engagement with a web-based food and exercise diary in a commercial weight loss programme: a retrospective analysis. *Int J Behav Nutr Phys Act* 2011;8:83 [[FREE Full text](#)] [doi: [10.1186/1479-5868-8-83](#)] [Medline: [21810222](#)]
55. Carlbring P, Brunt S, Bohman S, Austin D, Richards J, Öst L, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior* 2007 May;23(3):1421-1434. [doi: [10.1016/j.chb.2005.05.002](#)]
56. Bishop FL, Lewis G, Harris S, McKay N, Prentice P, Thiel H, et al. A within-subjects trial to test the equivalence of online and paper outcome measures: the Roland Morris disability questionnaire. *BMC Musculoskelet Disord* 2010;11:113 [[FREE Full text](#)] [doi: [10.1186/1471-2474-11-113](#)] [Medline: [20529332](#)]
57. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [[FREE Full text](#)] [doi: [10.2196/jmir.1923](#)] [Medline: [22209829](#)]

Abbreviations

- MOS:** Medical Outcomes Study
- RDQ:** Roland Morris Disability Questionnaire
- SD:** standard deviation
- VA:** Department of Veterans Affairs

Edited by G Eysenbach; submitted 07.03.13; peer-reviewed by M Buhrman, P Janwantanakul; comments to author 12.05.13; revised version received 03.06.13; accepted 18.06.13; published 22.08.13.

Please cite as:

Krein SL, Kadri R, Hughes M, Kerr EA, Piette JD, Holleman R, Kim HM, Richardson CR

Pedometer-Based Internet-Mediated Intervention For Adults With Chronic Low Back Pain: Randomized Controlled Trial

J Med Internet Res 2013;15(8):e181

URL: <http://www.jmir.org/2013/8/e181/>

doi: [10.2196/jmir.2605](https://doi.org/10.2196/jmir.2605)

PMID: [23969029](https://pubmed.ncbi.nlm.nih.gov/23969029/)

©Sarah L Krein, Reema Kadri, Maria Hughes, Eve A Kerr, John D Piette, Rob Holleman, Hyungjin Myra Kim, Caroline R Richardson. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 22.08.2013. 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 a Web 2.0 Portal to Improve Education and Communication in Young Patients With Families: Randomized Controlled Trial

Lena Hanberger¹, RN, PhD; Johnny Ludvigsson¹, MD, PhD; Sam Nordfeldt², MD, PhD

¹Division of Pediatrics, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

²Division of Child and Adolescent Psychiatry, Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

Corresponding Author:

Lena Hanberger, RN, PhD

Division of Pediatrics

Department of Clinical and Experimental Medicine

Linköping University

Linköping University Hospital

Linköping, 581 85

Sweden

Phone: 46 10 1031378

Fax: 46 13 148265

Email: lena.hanberger@lio.se

Abstract

Background: Diabetes requires extensive self-care and comprehensive knowledge, making patient education central to diabetes self-management. Web 2.0 systems have great potential to enhance health information and open new ways for patients and practitioners to communicate.

Objective: To develop a Web portal designed to facilitate self-management, including diabetes-related information and social networking functions, and to study its use and effects in pediatric patients with diabetes.

Methods: A Web 2.0 portal was developed in collaboration with patients, parents, and practitioners. It offered communication with local practitioners, interaction with peers, and access to relevant information and services. Children and adolescents with diabetes in a geographic population of two pediatric clinics in Sweden were randomized to a group receiving passwords for access to the portal or a control group with no access (n=230) for 1 year. All subjects had access during a second study year. Users' activity was logged by site and page visits. Health-related quality of life (HRQOL), empowerment (DES), and quality of information (QPP) questionnaires were given at baseline and after 1 and 2 study years. Clinical data came from the Swedish pediatric diabetes quality registry SWEDIABKIDS.

Results: There was a continuous flow of site visits, decreasing in summer and Christmas periods. In 119/233 families (51%), someone visited the portal the first study year and 169/484 (35%) the second study year. The outcome variables did not differ between intervention and control group. No adverse treatment or self-care effects were identified. A higher proportion of mothers compared to fathers visited once or more the first year ($P<.001$) and the second year ($P<.001$). The patients who had someone in the family visiting the portal 5 times or more, had shorter diabetes duration ($P=.006$), were younger ($P=.008$), had lower HbA1c after 1 year of access ($P=.010$), and were more often girls ($P<.001$). Peer interaction seems to be a valued aspect.

Conclusions: The Web 2.0 portal may be useful as a complement to traditional care for this target group. Widespread use of a portal would need integration in routine care and promotion by diabetes team members.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN):92107365; <http://www.controlled-trials.com/ISRCTN92107365/> (Archived by WebCite at <http://webcitation.org/6IkiIvtSb>).

(*J Med Internet Res* 2013;15(8):e175) doi:[10.2196/jmir.2425](https://doi.org/10.2196/jmir.2425)

KEYWORDS

type 1 diabetes; children; adolescent; health information technology; patient education; intervention studies; HbA1c

Introduction

Diabetes requires extensive self-care and comprehensive knowledge. The management of the disease, including insulin injections and self-control of blood glucose, affects everyday life, thus coping skills are essential. Health-related quality of life (HRQOL) may be influenced, particularly diabetes-related influence on HRQOL [1-3]. The association between good metabolic control and risk reduction for late complications is known [4-6] but despite modern treatment, only one third of the patients reach treatment target [7,8]. Efforts to increase patients' and parents' knowledge are needed to empower them in their self-care [9].

Thus patient education is central to diabetes self-management [10]. Studies in adult type 1 diabetes populations have indicated that structured patient training and education as part of intensive treatment reduces HbA1c with no increase in severe hypoglycemia, or even with persistent reduction of severe hypoglycemia [11-14]. Although such findings are consistent with modern clinical practice and experience [15], evidence repeatedly has been found insufficient to recommend adaptation of any particular educational method or program for type 1 diabetes [16,17]. There are several approaches, but there is no single one that emerges as clearly dominant.

We previously found that in pediatric patients' and parents' perspective on quality of care, improvements are needed regarding information and access to services [18]. In a multinational study, receiving information at diagnosis and having access to multiple sources of information later on have been associated with better outcomes from patients' and parents' perspectives [19]. The most frequently used sources of information, both for young adults and parents with diabetes, were diabetes medical teams, websites, and diabetes associations, with the diabetes team being the main source.

Social support is important for psychosocial adaptation when living with a pediatric chronic disease [20]. Recent research demonstrates how online support groups may contribute to patient empowerment [21]. Empowering processes identified among adult users of online support groups include exchanging information, encountering emotional support, finding recognition, sharing experiences, helping others, and amusement [21,22].

The Internet is a rapidly emerging source of health services and information [23]. Most adolescents and young adults find it convenient to use the Internet to communicate and find information, including on health topics [24], but still students lack knowledge about searching and evaluating health information on the Internet [25]. The umbrella term "Web 2.0" describes a range of widely used Internet applications to enhance participation, collaboration, openness, social networking, and peers' sharing information [26]. Web 2.0 systems have great potential to enhance health information delivery and exchange whenever and wherever it is needed, including use of new mobile devices.

Use of new information and communication technologies show promise regarding improved diabetes care in general [27-32].

At least those patients with poor metabolic control, greater use of health care services, higher motivation and/or less experience with diabetes treatment seem to benefit, although few significant long-term effects on main outcomes have been shown. Positive effects on knowledge and psychosocial well-being have been found as a result of Internet educational interventions in adolescents with diabetes [33].

Adolescents with a chronic health problem have been found to respond positively to sites that target them and their needs, including focused chat rooms and message boards [34]. In a pilot study, an Internet-based system aimed at assisting in diabetes management was found to be feasible and well accepted but did not influence HRQOL or metabolic control [35]. In a related study, communication with peers seemed to improve much more than their communication with practitioners [36]. Adolescents with diabetes visited various online forums for social support, information, advice, and shared experience [37]. Females used discussion forums more frequently and males requested more information.

In our study, we hypothesized that a Web 2.0 portal, with diabetes-related information and the possibility to communicate with diabetes peers as well as with health care professionals, would (1) be used, (2) be of complementary value in everyday life with diabetes, especially by newly diagnosed patients and patients in periods with instable metabolic control, (3) be perceived as helpful in self-treatment, and (4) contribute to improved metabolic control.

Thus we aimed to develop a Web portal designed to facilitate self-management, with diabetes-related information and the possibility to communicate with others with diabetes and health care professionals, and to study the use and its effects in pediatric patients with diabetes and their parents.

Methods

Ethics Statement

The study was approved by the Research Ethics Committee of the Faculty of Health Science at Linköping University, Sweden. Basic information about the study was given to adolescents and parents by posted letters. They were informed in the letter about confidentiality and the right to withdraw without explanation. All participants including next of kin were required to return a signed consent form. Informed consent was also given by each participant in electronic form prior to the first visit to the portal.

This study was a randomized controlled trial (ISRCTN92107365) and the CONSORT checklist is available as supporting information (see [Multimedia Appendix 1](#)).

The Portal

The Diabit Web 2.0 portal, as described elsewhere [38], offered self-directed communication with health professionals, interaction with peers, and access to information. The portal had been developed through a user-centred design process that included iterative sessions with groups of patients and parents as well as with the involved diabetes teams [39,40]. A prototype was piloted in 2005, and the portal Diabit was launched in April 2006. The portal was designed for complementary use by

pediatric patients, parents, and practitioners whenever needed and by users' own initiative.

It contained specific diabetes-related information and social networking functions such as a storyboard, a simple blog module, and discussion board modules (see examples of user interface in [Figures 1-3](#)). The discussion board in this version of the portal was designed for peer communication only (with safety issues monitored by a passive pediatrician). The practitioners involved did not have access to the patients' discussion board, and parents had no access to adolescents' board and vice versa.

Extensive information was given in text pages on essential areas of diabetes and in videos, and there was interactive simulator software as well [41]. Specific diabetes-related information on 13 main topics, developed by sessions with patients, and divided into 99 subtopics/Web pages, was written by team members. Links on diabetes-related information were: Acute situations,

What is diabetes, Relations, Late complications, Insulin, Devices, Food, Blood glucose, Exercise and sports, Living with diabetes, This can affect, Research, and External links. Each section was revised by other team members from the two hospitals.

The portal also provided services for medical prescription renewal, appointments, and open questions and other general information about the local diabetes teams and their services. In addition, each respective group of professionals comprising the two local diabetes teams summarized important basic information using a personal tone when expressing, "What I may say to newly diagnosed children and their parents".

To stimulate new visits, there was a function to highlight local practitioners' news and information about local activities, and for new information in the areas of research, nutrition, devices, and others. Quarterly newsletters were sent linking to the news, and flyers were sent yearly with regular post to patients.

Figure 1. Screenshot from the Diabit portal. Home "Welcome to Diabit".

diabit

Hem Om Diabit Regler Filmer Simulator Diskussionsforum Berättelser Dagbok Ordlista Blog Frågor A+ A- RESET

Akuta situationer
 Vad är diabetes?
 Vad händer i kroppen?
 Relationer
 Sena komplikationer
 Blodsocker
 Insulin
 Hjälpmedel
 Mat
 Motion och idrott
 Leva med diabetes
 Detta kan påverka
 Forskning
 Länkar

Sökord
 Sök

Oktoberfrågan: Hur ofta får du/ditt barn höra hårda/retfulla ord om diabetes?

- Aldrig hänt
- Enstaka tillfällen
- Flera tillfällen
- Ärligen
- Händer varje månad

Välkommen till Diabit!

- Till vänster finns info efter ämnen.
- Till höger diabetesteamet: maila dem, förnya recept, ända tider.
- Ovanför kan du se film, diskutera, blogga, simulera... testa själv!

Diabit - för ett bra liv med diabetes!

Diabits diskussionsforum
 Prata med andra... De som gör flest inlägg under okt vinner biobiljetter!
 Besök [Diabit forum här >>](#)
 Logga in som på Diabit
 Senast uppdaterad (2007-10-05)

Omega-3 tycks kunna minska risken för barndiabetes, enligt en amerikansk undersökning.
[Läs mer under forskningsnyheter >>](#)
 Senast uppdaterad (2007-09-28)

Nytt i Jönköping
 Barndiabetesmottagningen har flyttat till plan 6, hiss T [Läs mer >>](#)
 Senast uppdaterad (2007-09-20)

Nytt i Linköping
 Välkommen till vårt nästa kvällsmöte 9 oktober! [Läs mer >>](#)
 Senast uppdaterad (2007-08-31)

Tips om försäkringar
 För människor med diabetes kan det bli knepigt med försäkringar på flera sätt. Det kan dock löna sig att vara

Ny på Diabit?
 Gå till [Dina uppgifter](#) och skapa eget användarnamn och lösenord!
 Senast uppdaterad (2006-12-07)

Diabetesteamet Jönköping
Diabetesteamet Linköping
 För skolpersonal
 Logga ut
 Dina uppgifter

Figure 2. Screenshot from the Diabit portal. "I made a mistake with the insulin", and how to manage.

Akuta situationer

- Telefonnummer
- **Jag gjorde fel med insulinet!**
- Känningar
- Akutbehandling av känningar
- Ketoacidosis
- Akutbehandling av ketoacidosis

Vad är diabetes?

Vad händer i kroppen?

Relationer

Sena komplikationer

Blodsocker

Insulin

Hjälpmedel

Mat

Motion och idrott

Leva med diabetes

Detta kan påverka

Forskning

Länkar

Säkerhet

Jag gjorde fel med insulinet!

Om snabbinsulin gavs istället för basinsulin
Ät mat och följ blodsockeret en gång i timmen de närmaste fyra timmarna. Ge även halva dosen basinsulin direkt.

Om basinsulin gavs istället för snabbinsulin
Lägg till en halv dos snabbinsulin och minska ev. till halva dosen vid nästa måltid. Följ blodsockeret varannan timme i 6-8 timmar.
Om felet görs vid frukost eller lunch, ge vanliga kvälldosen basinsulin. Om felet görs vid middag eller senare, ge lägre dos basinsulin på kvällen.

Om rätt insulin gavs dubbelt
Ät motsvarande mer mat och följ blodsockeret en gång i timmen de närmaste timmarna. Hur länge beror på insulinsort. Vid direktverkande insulin följ blodsockeret varje halvtimme i början.

Skrivet av: Lena Hanberger, sjuksköt, Barnklin, US Linköping
Granskat av: Rosita Ilvered, barnsjuksköt, Britt-Marie Weidby, sjuksköt, Barnklin, Ryhov, Jönköping, Ann-Marie Sandström, barnsjuksköt och Ulf Samuelsson, barnläk, docent, Barnklin, US, Linköping
Senast uppdaterad (2007-08-24)

Diabetesteamet Jönköping

Diabetesteamet Linköping

För skolpersonal

Logga ut

Dina uppgifter

Figure 3. Screenshot from the Diabit portal. Discussion board.

> Visa olästa inlägg sen ditt senaste besök.
> Visa nya svar på dina inlägg.

Diabit > Föräldrar till barn med diabetes > Skolbarnsföräldrar > Motivation

Sidor: [1] SVARA | UNDERRÄTTA | MARKERA SOM OLÄST | SKICKA ÄMNET | SKRIV UT

Författare: JK_Vuxen
Amne: Motivation (läst 7 gånger)

Motivation
« skrivet: Oktober 09, 2007, 09:55:10 »

JK_Vuxen
Newbie
Antal inlägg: 3

Citat Ändra Radera Dela upp ämne

Vi har haft diabetes i vår familj i drygt 2 år. I början gick allt bra. Tills för ett halvår sedan. Vi visste att det skulle komma en reaktion, men när den kom var vi inte riktigt beredda på det. Nu går det bättre efter hjälp och lite samtal. Men det är fortfarande svårt med motivationen till att ta blodsocker samt insulin. Känns det igen? Kanske spelar åldern in (11 år)?
Berätta gärna om era erfarenheter.

Anmäl till moderator 82.182.16.248 (?)

Sidor: [1] SVARA | UNDERRÄTTA | LAGG TILL OMRÖSTNING | SKICKA ÄMNET | SKRIV UT

FLYTTA ÄMNE | RADERA ÄMNE | LÄS ÄMNE | KLISTRAS FAST ÄMNET | SAMMANFOGA ÄMNE

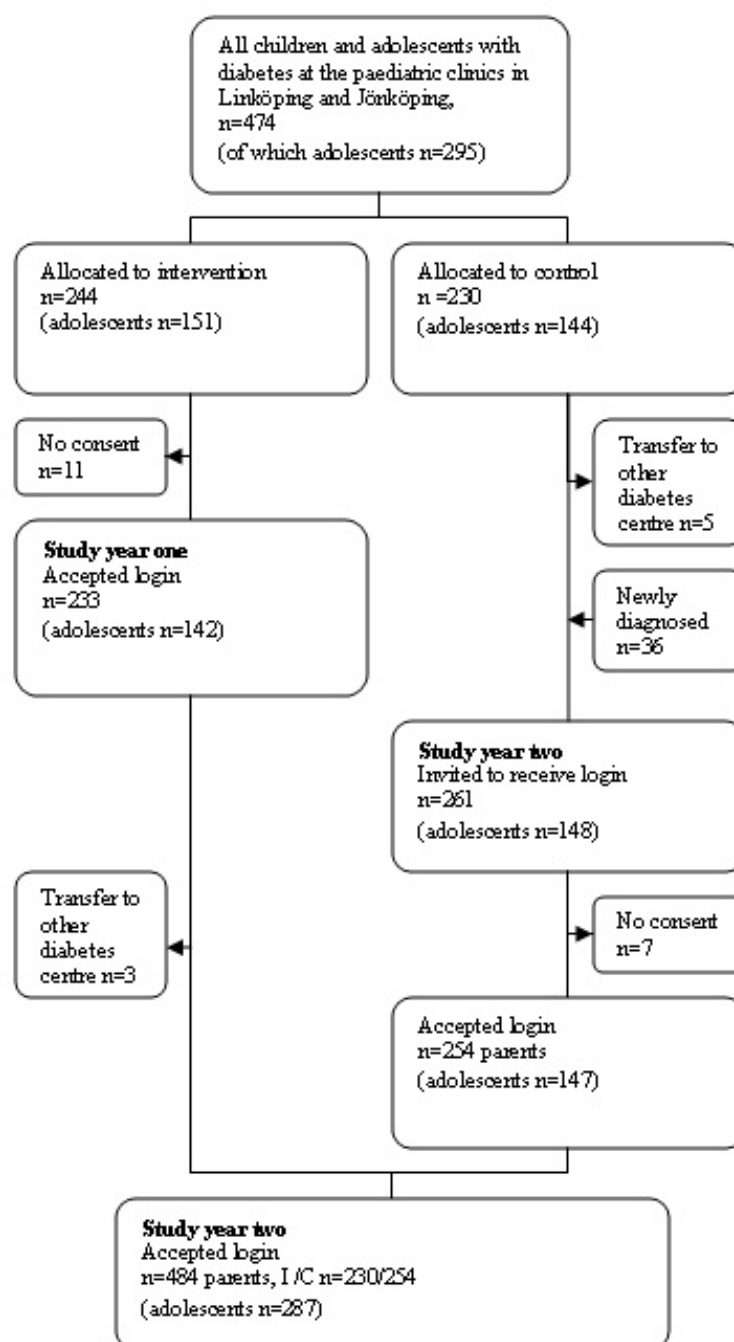
Protocol and Assignment

For inclusion, all the clinically diagnosed type 1 diabetes children, aged 0-18 years, registered in the Swedish pediatric diabetes quality registry, SWEDIABKIDS, belonging to the geographic population of the two pediatric clinics in Linköping and Jönköping, were eligible and invited to the study (Figure 4). The two clinics treated all young type 1 diabetes patients in their catchment areas. The patients and their families were randomized (stratified for clinic) by two of the authors (SN, LH), using a table of random numbers, to either the intervention group or the control group (Figure 4).

At baseline April 2006, all subjects in the intervention group were offered a personal password to the portal for the first year

of the study. After study year 1, all subjects in the previous control group were also offered passwords to the portal (Figure 4). For children 13 years of age and older, both parents and adolescents received passwords while for younger children, only parents received passwords.

As shown in Figure 4, for the first study year, 233 patients and their parents (adolescents $n=142$) accepted, and in the second study year, an additional 254 patients and their parents (adolescents $n=147$) from the previous control group accepted as well. All diabetes team members of both hospitals ($n=28$) received a personal password as well at baseline. During the study, there were no directions of use given to patients and parents from any other part, and it was not related to any structured education activity.

Figure 4. Flow chart of the intervention and the control group.

Masking

All personnel were blinded to group assignment for the first study year. Care providers took part in development of the portal, as described above, and were informed about the study but had no information on assignment. They were instructed to discuss any clinical issue raised by the patient as usual, without trying to identify the group to which the patient belonged.

Clinical Characteristics and Diabetes Treatment

The groups were equal regarding baseline clinical characteristics (Table 1). In Sweden, diabetes teams at pediatric clinics, consisting of diabetes specialist nurses, registered nurses, diabetes specialist physicians, dieticians, social workers and/or clinical psychologists, treat all patients in the catchment area. The treatment policy was multiple insulin therapy. The member of the diabetes team encouraged active self-control and offered psychosocial support and problem-based education. The process of care and treatment policy has been described elsewhere [15,42].

Table 1. Characteristics of the population at baseline, intervention and control group, and most recent HbA1c at baseline (2006) and after of study year 1 (2007).

	Intervention group n=244	Control group n=230	<i>P</i> value
Sex (F/M), proportions, %	52/48	51/49	.82
Mean age, years (range, SD)	13.2 (2.8-18.5, 3.7)	13.3 (3.0-18.5, 3.7)	.66
Age at diagnosis, years, mean (range, SD)	8.1 (1-17, 3.7)	7.8 (1-16, 4.2)	.50
Duration, years, mean (range, SD)	4.9 (0.1-17.7, 3.7)	5.1 (0.1-14.6, 3.7)	.67
HbA1c baseline, %, mean (range, SD)	6.8 (4.2-11.4, 1.2)	6.8 (4.3-12.5, 1.2)	.91
HbA1c after study year 1, %, mean (range, SD)	6.7 (4.1-13.0, 1.2)	6.7 (3.7-10.8, 1.2)	.72
Insulin dose, U/kg/day, mean (range, SD)	1.0 (0.2-2.1, 0.3)	1.0 (0.3-2.0, 0.3)	.95
CSII ^a , proportion, %	16	16	1.00
# of injections/day, mean (range, SD)	5.3 (2-9, 1.0)	5.4 (2-10, 1.0)	.37
# of self controls/day, mean (range, SD)	4.4 (0.5-12.5, 2.3)	4.1 (0-10, 1.7)	.51
Hypo last 12 months, proportion, %			
Needing help	16	25	.12
Unconsciousness	4	4	.71
Comorbidity	18	20	.16

^aCSII: continuous subcutaneous insulin infusion.

Study Period

The first study period, year 1, was defined as April 11, 2006, to September 25, 2007. The planned 12-month study period was extended due to initially slow inclusion of active users. The second study period, year 2, was defined as September 26, 2007, to September 25, 2008. However, there is a lack of logged data from August 30, 2008, due to problems with the data server.

Process Data

Logged data from the systems server were used to study frequencies and temporal patterns of patients and parents, as well as their practitioners, site visits, and page hits of the portal.

Outcome Variables

Effects on HRQOL, empowerment, and perception of quality of care regarding information were measured and obtained from postal surveys. The clinical variables measured were HbA1c (data received from the Swedish pediatric diabetes quality registry, SWEDIABKIDS) [43], numbers of severe hypoglycemia (self-reported), and numbers of self-controls of blood glucose (self-reported).

Questionnaires

For HRQOL, we used the DISABKIDS chronic-generic module, short form (12 items), adolescent, and parent (as proxy) version combined with the diabetes-specific module (10 items), adolescent and parent (as proxy) version [44,45]. The items in the chronic-generic module were assigned to six dimensions: independence, emotion, social inclusion, social exclusion, limitation, and medication. The items in the diabetes-specific module were assigned to treatment and impact on a 5-point Likert scale, where a low value corresponds to low quality of life.

Quality of care, regarding information, was measured by using selected questions from the Quality from the Patients' Perspective (QPP) questionnaire [46,47]. The QPP instrument was developed using grounded theory. The items evaluate both perceived reality of the care received and for subjective importance of that particular item, for example, "I get sufficient information regarding insulin pen/pump. A. This is how it is for me. B. This is how important it is for me." A 4-point scale ranging from "Fully agree" to "Do not agree at all" is used.

Empowerment was assessed by the Swedish Diabetes Empowerment Scale, short version (SWE-DES-SF-10) [48]. It includes four empowerment subscales: goal achievement, self-awareness, stress management, and readiness to change. A 5-point Likert scale is used.

Questions about access and using habits of the Internet, also used by Statistics Sweden [49] were included as well as a range of treatment-related questions and questions on socioeconomy, frequency of contact with peers, as well as online diabetes information search experiences.

Adolescents and parents completed questionnaires before baseline, posted late January 2006 (243 girls, 231 boys); after study year 1, posted late August 2007 (253 girls, 241 boys); and year 2, posted late August 2008 (250 girls, 234 boys) respectively. A mailed questionnaire and a stamped return envelope, with two subsequent reminders, were sent to all parents from an independent department at Linköping University. The response rates of the questionnaires were, in parents and adolescents respectively, at baseline 70% and 63%, after study year 1, 62% and 50%, and after study year 2, 59% and 65%.

Clinical Variables

Data from the Swedish pediatric diabetes quality registry, SWEDIABKIDS [35] were used regarding diabetes duration, hemoglobin A1c (HbA1c), and insulin dose. Swedish HbA1c values were approximately 1% lower than DCCT/National Glycohemoglobin Standardization Program (NGSP) values [50]. Local routine methods for HbA1c determination, calibrated to the national standard method Mono-S, were utilized.

Analysis

As a randomized control study, we compared the intervention and the control group at baseline and after study year 1. Additionally in both groups separately, baseline data were compared to data after study years 1 and 2. Most recent HbA1c values for each patient at baseline, at the end of study year 1 and at the end of study year 2 were used.

Active User Analysis

In a separate analysis before and after the first year of access, active users were defined as those where someone in the family logged in five times or more during their first year with access to the portal. This cut-off level for active use was defined retrospectively taking into account the distribution of frequency of use.

The group of active users were compared to those with zero to four site visits during the same time period. Thus we merged data for the intervention group at baseline and after 1 year only (study year 1), and for the previous control group before and after 1 year of access respectively (study year 2).

Statistical Methods

Summing the raw scores of the items in DISABKIDS representing each domain and dividing by the answers in a domain (at least 5/6 or 4/5 answers in each domain are required) resulted in mean domain scores. Grand mean of generic and diabetes-specific HRQOL was derived from summing the item mean score and dividing by the numbers of items. The scale for generic and diabetes-specific HRQOL was converted to a scale of 0-100, where 0 corresponds to 1 on the 5-point scale and 100 corresponds to 5. As primary endpoints for HRQOL, we used the mean of generic and diabetes-specific HRQOL and the mean of the dimensions within these. Total scale Swe-DES-SF-10 was calculated by summing the ten items and dividing by 10.

For comparisons, Mann-Whitney U test and Wilcoxon signed rank test were used and when data were normally distributed Student's *t* test, paired and unpaired was used. On categorical variables, Chi-square test was used. *P* values <.05 were regarded

as significant. Mean and SD are given. For statistical analysis, SPSS 17.0 software was used.

Results

Overview

No differences in the baseline characteristics of the population were found between the intervention and control group (Table 1). No differences was found regarding socioeconomics, access to the Internet, or information search and peer contacts.

Use of the Portal

During the very first month after launch, 51 users (14 adolescents, 26 mothers, 11 fathers) from 39 families of the intervention group visited the portal once or more (1456 page visits). The long-term pattern indicated a continuous interest for site visits, decreasing during summer and Christmas periods, as shown in Figure 5 (similar pattern for numbers of page visits and visitors, data not shown).

During the first study year, 159 users made 695 visits to the portal (adolescents 163, mothers 363 and fathers 169), mean 4.4 visits, range 1-45, median 2, and 6421 page hits (adolescents 1611, mothers 3484 and fathers 1326), mean 39.2, range 1-330, median 28.

During the second study year, 207 users made 980 visits (adolescents 210, mothers 573, and fathers 197), mean 4.7 visits, range 1-132, median 2, and 5940 page hits (adolescents 1954, mothers 3364, and fathers 622), mean 28.7, range 1-381, median 20. Thus the mean numbers of page visits per site visit in study year 1 was 9.2 (by adolescents 9.9, mothers 9.6, and fathers 7.8) and in study year 2, 6.1 (by adolescents 9.3, mothers 5.9, and fathers 3.0) respectively.

The proportions of those visiting the portal at least once or more during study year 1 and 2 respectively are shown in Figure 6, with higher proportions of mothers as compared to fathers the first ($P<.001$) and the second study year ($P<.001$). Out of those patients where someone in the family visited at least once during study year 1 ($n=119$, 51%) and year 2 ($n=169$, 35%) respectively, the proportions of active users (five times or more) were 30% the first study year and 64% the second study year.

More frequent page hits were seen during the first study year on social networking with peers such as Blogs and Stories followed by Questions answered by the diabetes team as well as their News and updates (Table 2). This pattern was largely similar during study year 2 (data not shown).

Table 2. Page hits on frequently visited pages, intervention group study year 1.

Webpages	Mothers n=71		Fathers n=39		Adolescents n=46	
	Hits	%	Hits	%	Hits	%
Home	358	10.4	168	12.7	163	10.1
Stories ^a	227	6.6	85	6.4	101	5.3
Blogs ^a	178	5.2	93	7.1	120	7.5
Team Jönköping	114	3.3	53	4.0	25	1.6
Questions and answers	109	3.1	17	1.3	29	1.8
Team Linköping	76	2.2	35	2.6	34	2.1
Research	75	2.2	38	2.9	15	.9
Simulator	62	1.8	36	2.7	37	2.3
Food	58	1.7	18	1.4		
This can affect	51	1.5	17	1.3	24	1.5
Devices	49	1.4	31	2.3		
Living with diabetes	43	1.2			26	1.6
Late complications	36	1.0	15	1.1	19	1.2
Videos	34	1.0	27	2.0	44	2.7
Discussion board ^a	34	1.0	12	.9	20	1.2
External links	32	.9				
Relations	30	.9	12	.9	20	1.2

^aSocial networking.

Figure 5. Site visits per month by patients and parents (study year 2 started 2007).

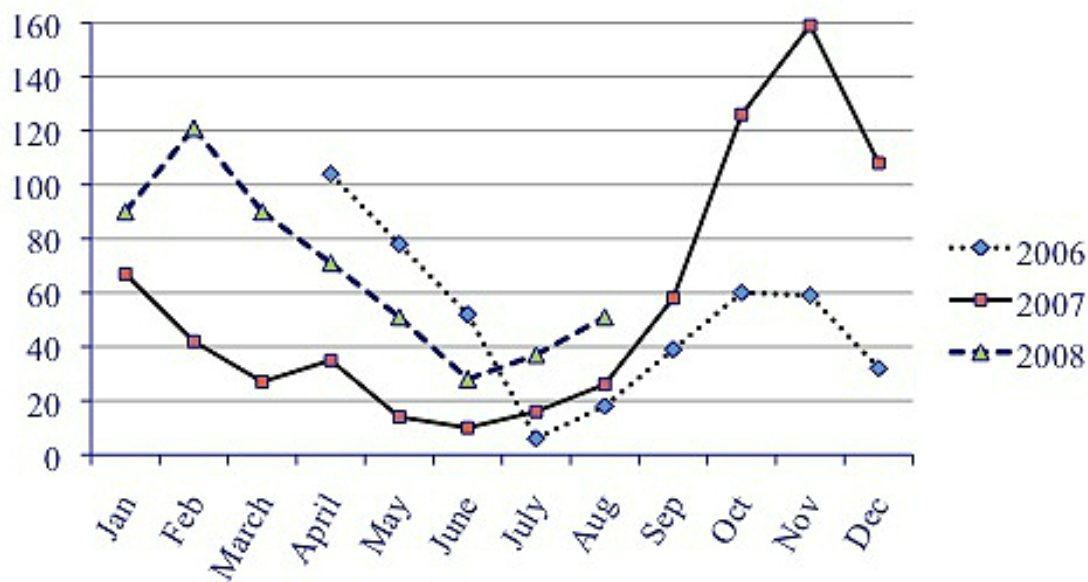
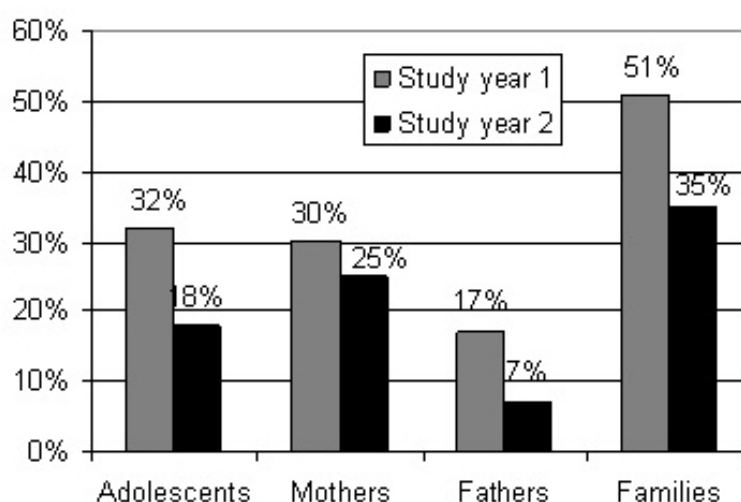


Figure 6. Proportions of adolescents, mothers, fathers, or at least one of these in the same family, logged in once or more.

The Intervention Group Compared to the Control Group

No differences were found at baseline and after study year 1 between the intervention and control group, adolescents and parents respectively, regarding the outcome variables (HRQOL, empowerment, perception of quality of care regarding information measured by DISABKIDS, SWE-DES-SF-10, and QPP respectively, HbA1c, severe hypoglycemia, frequency of blood glucose self-control). No differences were found at baseline and after study year 1 and 2 respectively, neither in the intervention group nor in the control group, for adolescents and parents respectively, regarding the same outcome variables.

Active Users

During the first year of access to the portal, active users (families where someone made five visits or more, $n=68$), compared to the less active (0-4 visits respectively, $n=419$) were younger,

had shorter durations, and had lower HbA1c after 1 year, but there was only a moderate—not significant—change in the HbA1c difference over time (Table 3). No differences were found among active users at 1 year compared to baseline regarding the questionnaires and clinical outcome variables. A higher proportion of girls' families were found among the active users ($P<.001$). A similar proportion of active users 7/68 (10%) had HbA1c >8 as in the comparison group, 56/419 (13%) (pediatric .535).

Practitioners' Use

A total of 20 users out of the 28 diabetes team members who received a password made 459 site visits during study year 1 and 2552 page visits, mean 5.6 page visits per site visit, range 3-370. In study year 2, they made 426 site visits and 1712 page visits, mean 4.0 page visits per site visit, range 6-826. Out of the 28 receiving passwords, 16 made ≥ 5 visits (15 users study year 1, 9 users study year 2).

Table 3. Active users (5 visits or more) as compared to less active (0-4 visits), including most recent HbA1c before access and at the end of 1 year of access.

	Active users	Comparison group	P value
Sex, % girls	65	49	<.001
Age, years, mean (range, SD)	12.1 (3.2-18.3, 3.8)	14.0 (1.9-18.5, 3.8)	.008
Duration, years, mean (range, SD)	3.9 (0.1-14.6, 3.3)	5.5 (0.1-17.7, 3.8)	.006
HbA1c before, %, mean (range, SD)	6.5 (3.9-9.0, 1.1)	6.8 (3.7 -11.4, 1.3)	.056
HbA1c at 1 year, %, mean (range, SD)	6.4 (4.2-10.10, 1.1)	6.8 (4.1-14.0, 1.2)	.010

Discussion

Principal Findings

The portal attracted all groups of users with great individual variation in frequency of use, and as expected, not all were users. In contrast to research on structured patient education including Web 2.0, this study did not evaluate a directed intervention program. The portal was a complementary information and communication resource for self-directed use whenever needed. Based on known information needs,

developed through patient, parent, and health practitioners' interaction, the portal merely offered a practitioner-driven high-quality alternative for online information and communication. Thus we performed this experiment as close to a future real-world resource as possible, leaving it to patients and parents to make their own decisions on use of the portal. In contrast to numerous other online resources though, the patients' local multiprofessional pediatric diabetes teams created and/or verified the information contents.

The active users' data seem somewhat promising, although the overall usage rate was low. As expected, a higher frequency of use was related to shorter diabetes experience. The frequency of use might also relate to factors such as personal interest and motivation, and/or perceived health status and satisfaction with the traditional care and education. For clinicians, it is a challenge to motivate and support young patients struggling with impaired metabolic control and this group must be given high priority for efforts. For patients reaching the HbA1c 8% level, interventions in our clinic are monthly visits, intensified adjustment of treatment, use of glucose sensors, and/or referral to the ward. As such patients were equally found among the active users, the portal appears useful also for this group of patients and their parents. This is in line with earlier findings that patients with poor metabolic control seem to benefit, sometimes even more than others, from the use of electronic communication [30].

In our qualitative evaluation presented elsewhere, users found the portal widely useful with the combination of peer support, reliable facts, and updates [38]. Being enabled to search when necessary and find reliable information provided by local clinicians was regarded as a great advantage, facilitating a feeling of security and being in control. Finding answers to difficult-to-ask questions, questions users did not know they had before, and questions focusing on sensitive areas were important. Visiting the portal could generate more information than expected, which could inspire increased use, and many had high positive expectations of a larger patient community. However, the password requirement appeared to be a real obstacle for some, probably limiting the number of users as well as frequency of use.

Already in the first study year, a high proportion of page hits on interactive functions indicated great interest in communication with peers and also for users' questions answered by practitioners (Table 2) consistent with other findings [51]. Adolescent users and parents as well submitted stories, blogs, and comments increasingly during both years. From the second study year, a lively discussion board was started by parents, whereas the adolescent discussion board remained largely silent both study years. The discussion board privacy concept with access only for the peer group involved did not prove advantageous, but many parents appreciated their board as such.

As reported elsewhere, although many participants submit few or no posts, the community attracts a high proportion of site visits and many derive passive support from reading about the experiences of others, both positive and negative [36,38]. Thus we believe an open access discussion board is more likely to create benefit for many, along with the development and impact of Web 2.0 towards openness [26]. Further, our data confirm previous findings that girls and mothers seem more active in searching for information and more eager to communicate electronically [32,37,52], and we also found a notable lower rate of boys' families among the active users. Further research in this area will need to be sensitive to gender differences.

Strengths and Limitations

Some strengths of this study are a large study base and the controlled design including randomization of patients to either the intervention or the control group. The development of the portal with parents, adolescents, and diabetes team members participating laid the foundation for an appealing portal and a user-friendly design. No undesired effects were found on self-perceived HRQOL, empowerment, and metabolic control, which happened in some intervention programs [53]. No incidents of any undesired treatment effects related to the portal use were reported from practitioners or patients.

The overall absence of statistically significant effects should be interpreted with caution because (1) the number of active users in the intervention group was low, (2) there was no promotion from the blinded practitioners in the randomized controlled phase, and (3) frequencies of use were limited by a password requirement having the function of a gatekeeper [38]. A limitation of the study is that the effect on patient and parent knowledge of diabetes was not evaluated.

Also during the second study year, in spite of positive attitudes over a long-term involvement in the development and contents [38,40], many practitioners had a hard time starting to make use of the portal in their practice (unpublished data). Various obstacles were reported, such as deep-rooted habits, having no computer in the room, or having too many working tasks. An important issue is how to increase engagement in patients and their next of kin. Whereas this study cannot answer this, practitioners' views recently have been further explored [54].

The Internet remains a rather new tool in patient education, and the implementation of using it is not a rapid process in routine care [38,40]. For practitioners trained in a culture of care with secrecy and strong restrictions regarding dissemination of patients' data, the global process towards openly sharing personal health information on the Internet [55] initially might seem somewhat uncomfortable and confusing. However, provided that practitioners can control and/or monitor the information content in the portal, over time their motivation for using it in their daily practice will probably increase.

During the second study year (second year of access for the intervention group), the overall proportion of the study participants who used the portal decreased, as some previous intervention group users presumably decreased their use after a longer time (Figure 6). Also in other interventions with Internet-based systems, a decline in use over time has been found [35,36]. In implementation of Web 2.0 systems for patients, attention should be paid to highlighting the feed of new information from their practitioners, as well as new messages and blogs posted from peer users, and strategies for external advertisements and reminders are needed as well [38].

To sum up, the logged user behaviors and our qualitative evaluations indicate that a fully implemented Web 2.0 system including a larger population for a community and without passwords might be of great complementary value for both patients and professionals [38]. Future research also involving larger sample sizes and with multicentre collaboration might add knowledge on development of various effective educational

interventions [17]. Patient engagement and social marketing of new tools warrant more attention.

Following this study in subsequent scientific experiments, the Diabit Web 2.0 portal was rebuilt and opened for free use on the open Internet, including an open discussion board, on World Diabetes Day November 14, 2008. The total user rate increased in 2009 to 29,015 yearly visits (144,336 page visits) from all over the nation, showing a continuous interest for both the growing community as well as factual information.

Conclusions

This study supports the fact that a Web 2.0 portal may be successfully used as a complement to traditional patient education and support. The implementation might be further enhanced by easy access without passwords, by highlighting new information, by active promotion from active diabetes team members and through other reminders in the structure of care. Future research on electronic communication targeting young people with long-term health problems will need to focus more on use of Web 2.0, including gender aspects.

Acknowledgments

We thank all the participating adolescents, parents, and diabetes team members in Jönköping and Linköping, who contributed views and ideas in developing the portal and performing the experiment.

Leni Ericson and Anders Larsson at IDA, Linköping University contributed to early design and prototype setup. Thor Balkhed, Bildamedia, Jörgen Isberg, Keylab, and Anders Larsson, IDA, Linköping University, contributed to elements of the portal design, setup and/or maintenance for the study.

The Linköping Diabit study was a practitioner-driven project supported by the Medical Research Council of Southeast Sweden (FORSS) and by ALF Grants, County Council of Östergötland, and the Swedish Child Diabetes Foundation (Barndiabetesfonden).

Authors' Contributions

All authors contributed to the design of the study. LH prepared the database and performed the analyses. All authors contributed to interpreting the findings, writing, or revising the manuscript, and approved the final version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [56].

[PDF File (Adobe PDF File), 1001KB - [jmir_v15i8e175_app1.pdf](#)]

References

1. Graue M, Wentzel-Larsen T, Hanestad BR, Båtsvik B, Søvik O. Measuring self-reported, health-related, quality of life in adolescents with type 1 diabetes using both generic and disease-specific instruments. *Acta Paediatr* 2003 Oct;92(10):1190-1196. [Medline: [14632337](#)]
2. Hanberger L, Ludvigsson J, Nordfeldt S. Health-related quality of life in intensively treated young patients with type 1 diabetes. *Pediatr Diabetes* 2009 Sep;10(6):374-381. [doi: [10.1111/j.1399-5448.2008.00496.x](#)] [Medline: [19207230](#)]
3. Hoey H, Aanstoot HJ, Chiarelli F, Daneman D, Danne T, Dorchy H, et al. Good metabolic control is associated with better quality of life in 2,101 adolescents with type 1 diabetes. *Diabetes Care* 2001 Nov;24(11):1923-1928. [Medline: [11679458](#)]
4. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993 Sep 30;329(14):977-986. [doi: [10.1056/NEJM199309303291401](#)] [Medline: [8366922](#)]
5. Nordwall M, Arnqvist HJ, Bojestig M, Ludvigsson J. Good glycemic control remains crucial in prevention of late diabetic complications--the Linköping Diabetes Complications Study. *Pediatr Diabetes* 2009 May;10(3):168-176. [doi: [10.1111/j.1399-5448.2008.00472.x](#)] [Medline: [19175900](#)]
6. Eeg-Olofsson K, Cederholm J, Nilsson PM, Zethelius B, Svensson AM, Gudbjörnsdóttir S, et al. Glycemic control and cardiovascular disease in 7,454 patients with type 1 diabetes: an observational study from the Swedish National Diabetes Register (NDR). *Diabetes Care* 2010 Jul;33(7):1640-1646 [FREE Full text] [doi: [10.2337/dc10-0398](#)] [Medline: [20424222](#)]
7. Danne T, Mortensen HB, Hougaard P, Lynggaard H, Aanstoot HJ, Chiarelli F, Hvidøre Study Group on Childhood Diabetes. Persistent differences among centers over 3 years in glycemic control and hypoglycemia in a study of 3,805 children and adolescents with type 1 diabetes from the Hvidøre Study Group. *Diabetes Care* 2001 Aug;24(8):1342-1347. [Medline: [11473067](#)]

8. Hanberger L, Samuelsson U, Lindblad B, Ludvigsson J, Swedish Childhood Diabetes Registry SWEDIABKIDS. A1C in children and adolescents with diabetes in relation to certain clinical parameters: the Swedish Childhood Diabetes Registry SWEDIABKIDS. *Diabetes Care* 2008 May;31(5):927-929. [doi: [10.2337/dc07-1863](https://doi.org/10.2337/dc07-1863)] [Medline: [18235047](https://pubmed.ncbi.nlm.nih.gov/18235047/)]
9. Funnell MM, Anderson RM. Patient empowerment: a look back, a look ahead. *Diabetes Educ* 2003;29(3):454-8, 460, 462 passim. [Medline: [12854337](https://pubmed.ncbi.nlm.nih.gov/12854337/)]
10. Swift PG. Diabetes education in children and adolescents. *Pediatr Diabetes* 2009 Sep;10 Suppl 12:51-57. [doi: [10.1111/j.1399-5448.2009.00570.x](https://doi.org/10.1111/j.1399-5448.2009.00570.x)] [Medline: [19754618](https://pubmed.ncbi.nlm.nih.gov/19754618/)]
11. Bott U, Bott S, Hemmann D, Berger M. Evaluation of a holistic treatment and teaching programme for patients with Type 1 diabetes who failed to achieve their therapeutic goals under intensified insulin therapy. *Diabet Med* 2000 Sep;17(9):635-643. [Medline: [11051282](https://pubmed.ncbi.nlm.nih.gov/11051282/)]
12. Plank J, Köhler G, Rakovac I, Semlitsch BM, Horvath K, Bock G, et al. Long-term evaluation of a structured outpatient education programme for intensified insulin therapy in patients with Type 1 diabetes: a 12-year follow-up. *Diabetologia* 2004 Aug;47(8):1370-1375. [doi: [10.1007/s00125-004-1456-x](https://doi.org/10.1007/s00125-004-1456-x)] [Medline: [15258736](https://pubmed.ncbi.nlm.nih.gov/15258736/)]
13. Sämman A, Mühlhauser I, Bender R, Hunger-Dathe W, Kloos C, Müller UA. Flexible intensive insulin therapy in adults with type 1 diabetes and high risk for severe hypoglycemia and diabetic ketoacidosis. *Diabetes Care* 2006 Oct;29(10):2196-2199. [doi: [10.2337/dc06-0751](https://doi.org/10.2337/dc06-0751)] [Medline: [17003292](https://pubmed.ncbi.nlm.nih.gov/17003292/)]
14. Schachinger H, Hegar K, Hermanns N, Straumann M, Keller U, Fehm-Wolfsdorf G, et al. Randomized controlled clinical trial of Blood Glucose Awareness Training (BGAT III) in Switzerland and Germany. *J Behav Med* 2005 Dec;28(6):587-594. [doi: [10.1007/s10865-005-9026-3](https://doi.org/10.1007/s10865-005-9026-3)] [Medline: [16222412](https://pubmed.ncbi.nlm.nih.gov/16222412/)]
15. Ludvigsson J, Bolli GB. Intensive insulin treatment in diabetic children. *Diabetes Nutr Metab* 2001 Oct;14(5):292-304. [Medline: [11806471](https://pubmed.ncbi.nlm.nih.gov/11806471/)]
16. Patient Education in Managing Diabetes (Patientutbildning vid diabetes). Stockholm: Swedish Council on Technology Assessment in Health Care; 2009. URL: <http://www.sbu.se/en/Published/Yellow/Patient-Education-in-Managing-Diabetes/> [accessed 2013-07-26] [WebCite Cache ID 6IPFE2pKA]
17. Murphy HR, Rayman G, Skinner TC. Psycho-educational interventions for children and young people with Type 1 diabetes. *Diabet Med* 2006 Sep;23(9):935-943. [doi: [10.1111/j.1464-5491.2006.01816.x](https://doi.org/10.1111/j.1464-5491.2006.01816.x)] [Medline: [16922699](https://pubmed.ncbi.nlm.nih.gov/16922699/)]
18. Hanberger L, Ludvigsson J, Nordfeldt S. Quality of care from the patient's perspective in pediatric diabetes care. *Diabetes Res Clin Pract* 2006 May;72(2):197-205. [doi: [10.1016/j.diabres.2005.10.009](https://doi.org/10.1016/j.diabres.2005.10.009)] [Medline: [16300853](https://pubmed.ncbi.nlm.nih.gov/16300853/)]
19. Peyrot M, International DAWN Youth Survey Group. The DAWN Youth WebTalk Study: methods, findings, and implications. *Pediatr Diabetes* 2009 Dec;10 Suppl 13:37-45. [doi: [10.1111/j.1399-5448.2009.00612.x](https://doi.org/10.1111/j.1399-5448.2009.00612.x)] [Medline: [19930225](https://pubmed.ncbi.nlm.nih.gov/19930225/)]
20. Nicholas DB, Fellner KD, Frank M, Small M, Hetherington R, Slater R, et al. Evaluation of an online education and support intervention for adolescents with diabetes. *Soc Work Health Care* 2012;51(9):815-827. [doi: [10.1080/00981389.2012.699507](https://doi.org/10.1080/00981389.2012.699507)] [Medline: [23078013](https://pubmed.ncbi.nlm.nih.gov/23078013/)]
21. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. *Qual Health Res* 2008 Mar;18(3):405-417. [doi: [10.1177/1049732307313429](https://doi.org/10.1177/1049732307313429)] [Medline: [18235163](https://pubmed.ncbi.nlm.nih.gov/18235163/)]
22. 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/)]
23. Eysenbach G. What is e-health? *J Med Internet Res* 2001;3(2):E20 [FREE Full text] [doi: [10.2196/jmir.3.2.e20](https://doi.org/10.2196/jmir.3.2.e20)] [Medline: [11720962](https://pubmed.ncbi.nlm.nih.gov/11720962/)]
24. Borzekowski DL, Rickert VI. Adolescent cybersurfing for health information: a new resource that crosses barriers. *Arch Pediatr Adolesc Med* 2001 Jul;155(7):813-817. [Medline: [11434849](https://pubmed.ncbi.nlm.nih.gov/11434849/)]
25. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. *J Med Internet Res* 2011;13(4):e102 [FREE Full text] [doi: [10.2196/jmir.1703](https://doi.org/10.2196/jmir.1703)] [Medline: [22155629](https://pubmed.ncbi.nlm.nih.gov/22155629/)]
26. 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](https://pubmed.ncbi.nlm.nih.gov/18725354/)]
27. Bellazzi R, Larizza C, Montani S, Riva A, Stefanelli M, d'Annunzio G, et al. A telemedicine support for diabetes management: the T-IDDM project. *Comput Methods Programs Biomed* 2002 Aug;69(2):147-161. [Medline: [12100794](https://pubmed.ncbi.nlm.nih.gov/12100794/)]
28. Jackson CL, Bolen S, Brancati FL, Batts-Turner ML, Gary TL. A systematic review of interactive computer-assisted technology in diabetes care. *Interactive information technology in diabetes care. J Gen Intern Med* 2006 Feb;21(2):105-110 [FREE Full text] [doi: [10.1111/j.1525-1497.2005.00310.x](https://doi.org/10.1111/j.1525-1497.2005.00310.x)] [Medline: [16390512](https://pubmed.ncbi.nlm.nih.gov/16390512/)]
29. Mulvaney SA, Rothman RL, Wallston KA, Lybarger C, Dietrich MS. An internet-based program to improve self-management in adolescents with type 1 diabetes. *Diabetes Care* 2010 Mar;33(3):602-604 [FREE Full text] [doi: [10.2337/dc09-1881](https://doi.org/10.2337/dc09-1881)] [Medline: [20032275](https://pubmed.ncbi.nlm.nih.gov/20032275/)]
30. Verhoeven F, van Gemert-Pijnen L, Dijkstra K, Nijland N, Seydel E, Steehouder M. The contribution of teleconsultation and videoconferencing to diabetes care: a systematic literature review. *J Med Internet Res* 2007;9(5):e37 [FREE Full text] [doi: [10.2196/jmir.9.5.e37](https://doi.org/10.2196/jmir.9.5.e37)] [Medline: [18093904](https://pubmed.ncbi.nlm.nih.gov/18093904/)]

31. 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/)]
32. Franklin VL, Greene A, Waller A, Greene SA, Pagliari C. Patients' engagement with "Sweet Talk" - a text messaging support system for young people with diabetes. *J Med Internet Res* 2008;10(2):e20 [FREE Full text] [doi: [10.2196/jmir.962](https://doi.org/10.2196/jmir.962)] [Medline: [18653444](https://pubmed.ncbi.nlm.nih.gov/18653444/)]
33. Cooper H, Cooper J, Milton B. Technology-based approaches to patient education for young people living with diabetes: a systematic literature review. *Pediatr Diabetes* 2009 Nov;10(7):474-483. [doi: [10.1111/j.1399-5448.2009.00509.x](https://doi.org/10.1111/j.1399-5448.2009.00509.x)] [Medline: [19490492](https://pubmed.ncbi.nlm.nih.gov/19490492/)]
34. Flicker S, Goldberg E, Read S, Veinot T, McClelland A, Saulnier P, et al. HIV-positive youth's perspectives on the Internet and e-health. *J Med Internet Res* 2004 Sep 29;6(3):e32 [FREE Full text] [doi: [10.2196/jmir.6.3.e32](https://doi.org/10.2196/jmir.6.3.e32)] [Medline: [15471758](https://pubmed.ncbi.nlm.nih.gov/15471758/)]
35. Powell J, Jennings A, Armstrong N, Sturt J, Dale J. Pilot study of a virtual diabetes clinic: satisfaction and usability. *J Telemed Telecare* 2009;15(3):150-152. [doi: [10.1258/jtt.2009.003014](https://doi.org/10.1258/jtt.2009.003014)] [Medline: [19364901](https://pubmed.ncbi.nlm.nih.gov/19364901/)]
36. Jennings A, Powell J, Armstrong N, Sturt J, Dale J. A virtual clinic for diabetes self-management: pilot study. *J Med Internet Res* 2009;11(1):e10 [FREE Full text] [doi: [10.2196/jmir.1111](https://doi.org/10.2196/jmir.1111)] [Medline: [21821504](https://pubmed.ncbi.nlm.nih.gov/21821504/)]
37. Ravert RD, Hancock MD, Ingersoll GM. Online forum messages posted by adolescents with type 1 diabetes. *Diabetes Educ* 2004;30(5):827-834. [Medline: [15510534](https://pubmed.ncbi.nlm.nih.gov/15510534/)]
38. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 Diabetes Portal--the management tool, the generator, and the gatekeeper: qualitative study. *J Med Internet Res* 2010;12(2):e17 [FREE Full text] [doi: [10.2196/jmir.1267](https://doi.org/10.2196/jmir.1267)] [Medline: [20511179](https://pubmed.ncbi.nlm.nih.gov/20511179/)]
39. Ekberg J, Ericson L, Timpka T, Eriksson H, Nordfeldt S, Hanberger L, et al. Web 2.0 systems supporting childhood chronic disease management: design guidelines based on information behaviour and social learning theories. *J Med Syst* 2010 Apr;34(2):107-117. [Medline: [20433049](https://pubmed.ncbi.nlm.nih.gov/20433049/)]
40. Nordqvist C, Hanberger L, Timpka T, Nordfeldt S. Health professionals' attitudes towards using a Web 2.0 portal for child and adolescent diabetes care: qualitative study. *J Med Internet Res* 2009;11(2):e12 [FREE Full text] [doi: [10.2196/jmir.1152](https://doi.org/10.2196/jmir.1152)] [Medline: [19403464](https://pubmed.ncbi.nlm.nih.gov/19403464/)]
41. Nordfeldt S, Hanberger L, Malm F, Ludvigsson J. Development of a PC-based diabetes simulator in collaboration with teenagers with type 1 diabetes. *Diabetes Technol Ther* 2007 Feb;9(1):17-25. [doi: [10.1089/dia.2006.0053](https://doi.org/10.1089/dia.2006.0053)] [Medline: [17316094](https://pubmed.ncbi.nlm.nih.gov/17316094/)]
42. Nordfeldt S, Ludvigsson J. Adverse events in intensively treated children and adolescents with type 1 diabetes. *Acta Paediatr* 1999 Nov;88(11):1184-1193. [Medline: [10591418](https://pubmed.ncbi.nlm.nih.gov/10591418/)]
43. The Swedish pediatric diabetes quality registry. URL: <https://www.ndr.nu/ndr2/Default.aspx> [accessed 2012-10-01] [WebCite Cache ID 6B64eSkN3]
44. Baars RM, Atherton CI, Koopman HM, Bullinger M, Power M, DISABKIDS group. The European DISABKIDS project: development of seven condition-specific modules to measure health related quality of life in children and adolescents. *Health Qual Life Outcomes* 2005;3:70 [FREE Full text] [doi: [10.1186/1477-7525-3-70](https://doi.org/10.1186/1477-7525-3-70)] [Medline: [16283947](https://pubmed.ncbi.nlm.nih.gov/16283947/)]
45. Bullinger M, Schmidt S, Petersen C, DISABKIDS Group. Assessing quality of life of children with chronic health conditions and disabilities: a European approach. *Int J Rehabil Res* 2002 Sep;25(3):197-206. [Medline: [12352173](https://pubmed.ncbi.nlm.nih.gov/12352173/)]
46. Larsson G, Larsson BW, Munck IM. Refinement of the questionnaire 'quality of care from the patient's perspective' using structural equation modelling. *Scand J Caring Sci* 1998;12(2):111-118. [Medline: [9801632](https://pubmed.ncbi.nlm.nih.gov/9801632/)]
47. Wilde B, Larsson G, Larsson M, Starrin B. Quality of care. Development of a patient-centred questionnaire based on a grounded theory model. *Scand J Caring Sci* 1994;8(1):39-48. [Medline: [8159889](https://pubmed.ncbi.nlm.nih.gov/8159889/)]
48. Leksell J, Funnell M, Sandberg G, Smide B, Wiklund G, Wikblad K. Psychometric properties of the Swedish Diabetes Empowerment Scale. *Scand J Caring Sci* 2007 Jun;21(2):247-252. [doi: [10.1111/j.1471-6712.2007.00463.x](https://doi.org/10.1111/j.1471-6712.2007.00463.x)] [Medline: [17559444](https://pubmed.ncbi.nlm.nih.gov/17559444/)]
49. Statistics Sweden. URL: <http://www.scb.se/> [accessed 2012-10-01] [WebCite Cache ID 6B65c8uxR]
50. Hoelzel W, Weykamp C, Jeppsson JO, Miedema K, Barr JR, Goodall I, IFCC Working Group on HbA1c Standardization. IFCC reference system for measurement of hemoglobin A1c in human blood and the national standardization schemes in the United States, Japan, and Sweden: a method-comparison study. *Clin Chem* 2004 Jan;50(1):166-174 [FREE Full text] [doi: [10.1373/clinchem.2003.024802](https://doi.org/10.1373/clinchem.2003.024802)] [Medline: [14709644](https://pubmed.ncbi.nlm.nih.gov/14709644/)]
51. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. *J Med Internet Res* 2011;13(3):e71 [FREE Full text] [doi: [10.2196/jmir.1603](https://doi.org/10.2196/jmir.1603)] [Medline: [21959968](https://pubmed.ncbi.nlm.nih.gov/21959968/)]
52. Suris JC, Akre C, Berchtold A, Bélanger RE, Michaud PA. Chronically connected? Internet use among adolescents with chronic conditions. *J Adolesc Health* 2010 Feb;46(2):200-202. [doi: [10.1016/j.jadohealth.2009.07.008](https://doi.org/10.1016/j.jadohealth.2009.07.008)] [Medline: [20113927](https://pubmed.ncbi.nlm.nih.gov/20113927/)]
53. Viklund G, Ortqvist E, Wikblad K. Assessment of an empowerment education programme. A randomized study in teenagers with diabetes. *Diabet Med* 2007 May;24(5):550-556. [doi: [10.1111/j.1464-5491.2007.02114.x](https://doi.org/10.1111/j.1464-5491.2007.02114.x)] [Medline: [17367306](https://pubmed.ncbi.nlm.nih.gov/17367306/)]
54. Nordfeldt S, Ängarne-Lindberg T, Berterö C. To use or not to use--practitioners' perceptions of an open web portal for young patients with diabetes. *J Med Internet Res* 2012;14(6):e154 [FREE Full text] [doi: [10.2196/jmir.1987](https://doi.org/10.2196/jmir.1987)] [Medline: [23137767](https://pubmed.ncbi.nlm.nih.gov/23137767/)]

55. 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](#)]
56. 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](#)]

Abbreviations

CSII: subcutaneous continuous insulin infusion

DCCT: Diabetes Control and Complications Trial Research Group

DES: diabetes empowerment scale

HbA1c: hemoglobin A1c

HRQOL: health-related quality of life

NGSP: National Glycohemoglobin Standardization Program

QPP: quality from the patients' perspective

SWEDIABKIDS: Swedish pediatric diabetes quality registry

SWE-DES-SF-10: Swedish diabetes empowerment scale, short version

Edited by G Eysenbach; submitted 08.11.12; peer-reviewed by S Mulvaney, A Mostaghimi, M Hendrix; comments to author 08.03.13; revised version received 19.04.13; accepted 12.06.13; published 23.08.13.

Please cite as:

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

URL: <http://www.jmir.org/2013/8/e175/>

doi: [10.2196/jmir.2425](https://doi.org/10.2196/jmir.2425)

PMID: [23973555](https://pubmed.ncbi.nlm.nih.gov/23973555/)

©Lena Hanberger, Johnny Ludvigsson, Sam Nordfeldt. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 23.08.2013. 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

Barriers and Facilitators for the Implementation of an Online Clinical Health Community in Addition to Usual Fertility Care: A Cross-Sectional Study

Johanna WM Aarts¹, MD, PhD; Marjan J Faber², PhD; Anne G den Boogert¹, MD; Ben J Cohlen³, MD, PhD; Paul JQ van der Linden⁴, MD, PhD; Jan AM Kremer¹, MD, PhD; Willianne LDM Nelen¹, MD, PhD

¹Radboud University Nijmegen Medical Center, Department of Obstetrics and Gynecology, Radboud University, Nijmegen, Netherlands

²Radboud University Nijmegen Medical Center, Scientific Institute for Quality of Healthcare, Nijmegen, Netherlands

³Isala Clinics, Fertility Center Isala, Zwolle, Netherlands

⁴Deventer Hospital, Department of Obstetrics and Gynecology, Deventer, Netherlands

Corresponding Author:

Johanna WM Aarts, MD, PhD

Radboud University Nijmegen Medical Center

Department of Obstetrics and Gynecology

Radboud University

PO Box 9101

Nijmegen, 6500 HB

Netherlands

Phone: 31 243668666

Fax: 31 243617351

Email: a.aarts@obgyn.umcn.nl

Abstract

Background: Online health communities are becoming more popular in health care. Patients and professionals can communicate with one another online, patients can find peer support, and professionals can use it as an additional information channel to their patients. However, the implementation of online health communities into daily practice is challenging. These challenges relate to the fact that patients need to be activated to (1) become a member (ie, subscription) and (2) participate actively within the community before any effect can be expected. Therefore, we aimed at answering 2 research questions: (1) what factors are associated with subscription to an online health community, and (2) which are associated with becoming an active participant within an online health community.

Objective: To identify barriers and facilitators as perceived by patients for the implementation of an online health community.

Methods: We performed a cross-sectional study. Three Dutch fertility clinics (2 IVF-licensed) offered their patients a secure online clinical health community through which clinicians can provide online information and patients can ask questions to the medical team or share experiences and find support from peers. We randomly selected and invited 278 men and women suffering from infertility and attending 1 of the participating clinics. Participants filled out a questionnaire about their background characteristics and current use of the online community. Possible barriers and facilitators were divided into 2 parts: (1) those for subscription to the community, and (2) those for active participation in the community. We performed 2 multivariate logistic regression analyses to calculate determinants for both subscription and active participation.

Results: Subscription appeared to be associated with patients' background characteristics (eg, gender, treatment phase), intervention-related facilitators (odds ratio [OR] 2.45, 95% CI 1.14-5.27), and patient-related barriers (OR 0.20, 95% CI 0.08-0.54), such as not feeling the need for such an online health community. After subscription, determinants for participation consisted of aspects related to participant's age (OR 0.86, 95% CI 0.76-0.97), length of infertility (OR 1.48, 95% CI 1.09-2.02), and to intervention-related facilitators (OR 5.79, 95% CI 2.40-13.98), such as its reliable character and possibility to interact with the medical team and peers.

Conclusions: Implementing an online health community in addition to usual fertility care should be performed stepwise. At least 2 strategies are needed to increase the proportion of patient subscribers and consequently make them active participants. First, the marketing strategy should contain information tailored to different subgroups of the patient population. Second, for a living online health community, incorporation of interactive elements, as well as frequent news and updates are needed. These

results imply that involving patients and their needs into the promotion strategy, community's design, and implementation are crucial.

(*J Med Internet Res* 2013;15(8):e163) doi:[10.2196/jmir.2098](https://doi.org/10.2196/jmir.2098)

KEYWORDS

community networks; infertility; Internet; quality of health care; patient-centered care

Introduction

In health care today, it is of pivotal importance to take into account the patient's perspective of care. Patients wish to play an active role, are informed, and prefer involvement in the decision-making process [1-4]. This societal trend is especially visible in the field of reproductive medicine. A plethora of studies have described the importance of involving the patient's perspective in fertility care and addressed the switch toward more collaboration and partnership with our patients [5-11]. Patients need support from peers, prefer complete and reliable information, wish to communicate online with their clinicians, and want to have easier access to care [12-14]. The developments around Web 2.0, in which the Internet is used as an interactive medium characterized by participation and collaboration between people on the Internet [15-16], provides us with possibilities to fulfill these patients' needs. Web 2.0 technologies can integrate large amounts of information, which is especially useful in the rapidly evolving field of reproductive medicine in which new insights come and go [17]. Moreover, the Internet can also connect patients to others who are facing the same problem more simply than clinicians can [18-20]. In this respect, the usage of Web 2.0 technologies, such as forums and blogs, are gaining a more prominent position within health care [18,21,22].

The use of these technologies in online health communities in addition to usual care is gaining popularity [18,23]. Previous studies indicated that the integration of Web 2.0 technologies in health care might bring benefits for both patients and professionals in terms of patient empowerment and the possibility to tailor care more appropriately to the needs of patients, also known as patient-centeredness of care [14,21,23-25]. Also, the increasing demand from patients for such communities have led several health care organizations, such as Johns Hopkins Hospital and The Cleveland Clinic, to establish online communities and discussion forums as part of their patient-support services [26]. However, adoption of online health communities is challenging and many interventions lack the ability to maintain usage in the long term [22,27-30]. Potential users should be tempted to join the online health community and, for sustainability, he or she also needs to be challenged to participate actively [30,31]. Chiu and Eysenbach [31] identified 4 stages of using Internet-based interventions that are relevant before positive outcomes can be expected: (1) consideration, (2) initiation, (3) utilization, and (4) outcomes. Every stage has its own barriers, of which adjustment might eventually improve the implementation. Thus, systematically inventorying these factors that facilitate or hinder the use of these interventions is crucial in developing targeted and effective implementation strategies [32].

In this cross-sectional study, we aimed at identifying the barriers and facilitators for the implementation of an online health community in addition to usual fertility care. Therefore, we aimed at answering 2 research questions: (1) what factors are associated with subscription to an online health community, and (2) which are associated with becoming an active participant within an online health community?

Methods

Setting

In the Netherlands, couples with impaired fertility can be referred by their general practitioner to a gynecologist in a hospital for further assessment of their fertility problem and for intrauterine insemination (IUI) and ovulation induction (OI) as the first treatment possibilities. In vitro fertilization (IVF), including intracytoplasmic sperm injection (ICSI), is only performed in 13 IVF-licensed clinics in the Netherlands. In some hospitals without an IVF laboratory, physicians can start up and monitor IVF, perform the oocyte retrieval, and then refer the patient to an IVF clinic for embryo transfer (transport clinic). The Dutch national health care system reimburses the costs of the diagnostic work up, 6 IUI and all OI cycles, and the first 3 IVF cycles. The clinics participating in this study were 2 IVF-licensed clinics and 1 transport clinic.

Description of an Online Health Community in Addition to Usual Fertility Care

An online health community was constructed as a members-only online community provided by an online platform for online health communities, *MijnZorgnet (MyCareNet)* [33]. An online health community offered several functions. First, by means of blogs, professionals could inform their patients about relevant news. Second, it provided 2 separate discussion forums: one in which patients could share experiences and communicate with one another, the other in which patients could ask questions to the medical team. Third, it contained a media gallery in which patients could find digital information leaflets on infertility-related topics. The 3 clinics participating in this study offered such a secured online health community to their own patient population in addition to usual care.

The setup of an online health community was initiated by the head of the department of the 3 different clinics and aimed for improvement of patient-centeredness of care. In every clinic, a nurse or medical assistant was assigned to act as the community manager, responsible for maintenance of the online health community. To become a member, patients used their personal digital identification code to create a profile on the platform of *MijnZorgnet* [33]. After log-in, patients had to send a membership request to get access. Patients were granted access after subscription with their patient identification number of the

hospital. At all 3 clinics, generic information leaflets about the online health community were distributed personally to invite infertile patients to become a member. These patients had their intake visit, underwent a diagnostic work up, or had a fertility treatment, including OI, IUI, or IVF/ICSI.

Development of Questionnaire

The questionnaire was aimed at identifying aspects relevant to subscribing and active participating in the online health communities. The first part of the questionnaire consisted of questions on background characteristics (eg, age) and characteristics related to their fertility problems (eg, treatment). The second part included items concerning possible barriers and facilitators for subscription to the online health community (part 1), and barriers and facilitators for active participation within the online health community (part 2). Items for this part of the questionnaire were generated from semistructured interviews with 8 patients, conducted for this purpose. All 8 patients had heard about the community, but only 6 decided to subscribe. These patients were asked about the aspects that may impede or facilitate subscription to and participation in the online community and its value for current health care. Interviews were recorded and transcribed verbatim. Transcripts were thematically analyzed by 2 researchers independently and discussed among them to increase coding reliability. Then they divided these items independently into possible barriers and possible facilitators for subscription and participation respectively. They used the 4 domains according the framework of Cabana et al [34] as a framework: patient-related characteristics, intervention-related characteristics, professional-related characteristics, and characteristics of the context in which the intervention was applied. Differences in categorization between researchers were small and consensus was mostly promptly achieved. Although we chose to base the internal consistency of these domains on rigorously performed qualitative analysis, we also calculated Cronbach alpha for each domain as additional information for readers.

These 46 items were converted to a statement. Patients answered at a 4-point Likert scale indicating total disagreement (1) to total agreement (4) with a particular item as a barrier or facilitator for subscribing to or participating in the online health community. All barriers and facilitators were applicable for both subscribing to and participating in the community. Others only applied to active participation, such as “the website doesn’t encourage posting comments or reactions.”

The final questionnaire was pretested among 5 patients resulting in few textual adjustments and the removal of 2 questions.

Participants and Data Collection

We invited patients who attended 1 of the 3 fertility clinics that participated in this study. We aimed at inviting both patients who were a member of the online health community and patients who were informed about the startup of the online infertility community, but did not subscribe to the community. From the online infertility communities’ members databases, the main researcher randomly selected half of the patients (n=141) to participate in the study. To identify patients who had not subscribed to the online infertility community, the community

managers listed all patients that visited the clinic in the previous 2 weeks for an intake consultation, diagnostic assessments, or a fertility treatment. We deleted patients from the lists who already subscribed to the online infertility community. Thereafter, we randomly selected patients from these lists and invited both partners of a couple separately to participate in this study. The proportion of subscribed versus nonsubscribed patients was 1:2, foreseeing a lower response rate of nonsubscribed patients. All participants received a questionnaire package by mail 6 months after the setup of the online infertility community. The questionnaire package was accompanied by instructions, a refusal form, and a stamped return envelope. Participation in the study was voluntary and anonymous. In the Netherlands, institutional ethics committee approval was not required for this study. Participants were sent a reminder at 3 and 5 weeks following the initial mailing, respectively. Figure 1 presents an overview of the data collection and analysis procedure.

Data Analysis

Overview

Data from incoming questionnaires were entered into SPSS version 16.0 for Windows (SPSS Inc, Chicago, IL, USA). Participants who filled out less than 50% of the questionnaire were removed from the database. We used descriptive statistics to present background characteristics of the study population. Answers to open-ended questions were synthesized and categorized. We performed bivariate and multivariate logistic regression analyses to determine factors associated with subscription to (analysis 1) and active participation in (analysis 2) the online infertility community.

Independent Variables

In both analyses, we used all patients’ background characteristics (part 1 of questionnaire) combined with the 7 categories of barriers and facilitators (eg, intervention-related category; see Table 1) as independent variables that were based on rigorously performed qualitative analysis. For analysis 1, we used the categories that were composed of those items that were only applicable for subscription (see Table 1). For analysis 2, we used all 7 categories, composed of the 44 single items. Table 1 also shows the statistical reliability of these categories presented as Cronbach alpha. For both analyses, we used per category mean sum scores calculated as the mean score of each individual item divided by the number of items within the category.

Dependent Variables

For analysis 1, the dichotomous dependent outcome variable included the question whether they subscribed or did not subscribe to the online infertility community (0=no; 1=yes). In analysis 2, the dependent variable consisted of the activity of a participant within the online infertility community (0 = inactive; 1 = active). We categorized the latter based on self-reported activity. Inactive members had not visited the online infertility community at all after subscription or just a few times without further action. Active users had read the content, visited the online infertility community daily, posted messages, or asked online questions to the medical team. These categories were derived from a social participation ladder [35].

In both analyses, we performed Pearson correlation tests to check for collinearity between the independent variables. Whenever a correlation between 2 variables was more than 0.6, we excluded 1 of those from further analysis. Then, we conducted bivariate logistic regression analysis for each of the independent variables with the 2 different dependent variables.

Variables with $P < .20$ were found to be eligible for multivariate regression analysis. A backward selection method was applied, and we considered factors with $P < .05$ significant. We calculated adjusted odds ratios (ORs), P values, and 95% confidence intervals (95% CI).

Figure 1. Overview of inclusion procedure participants.

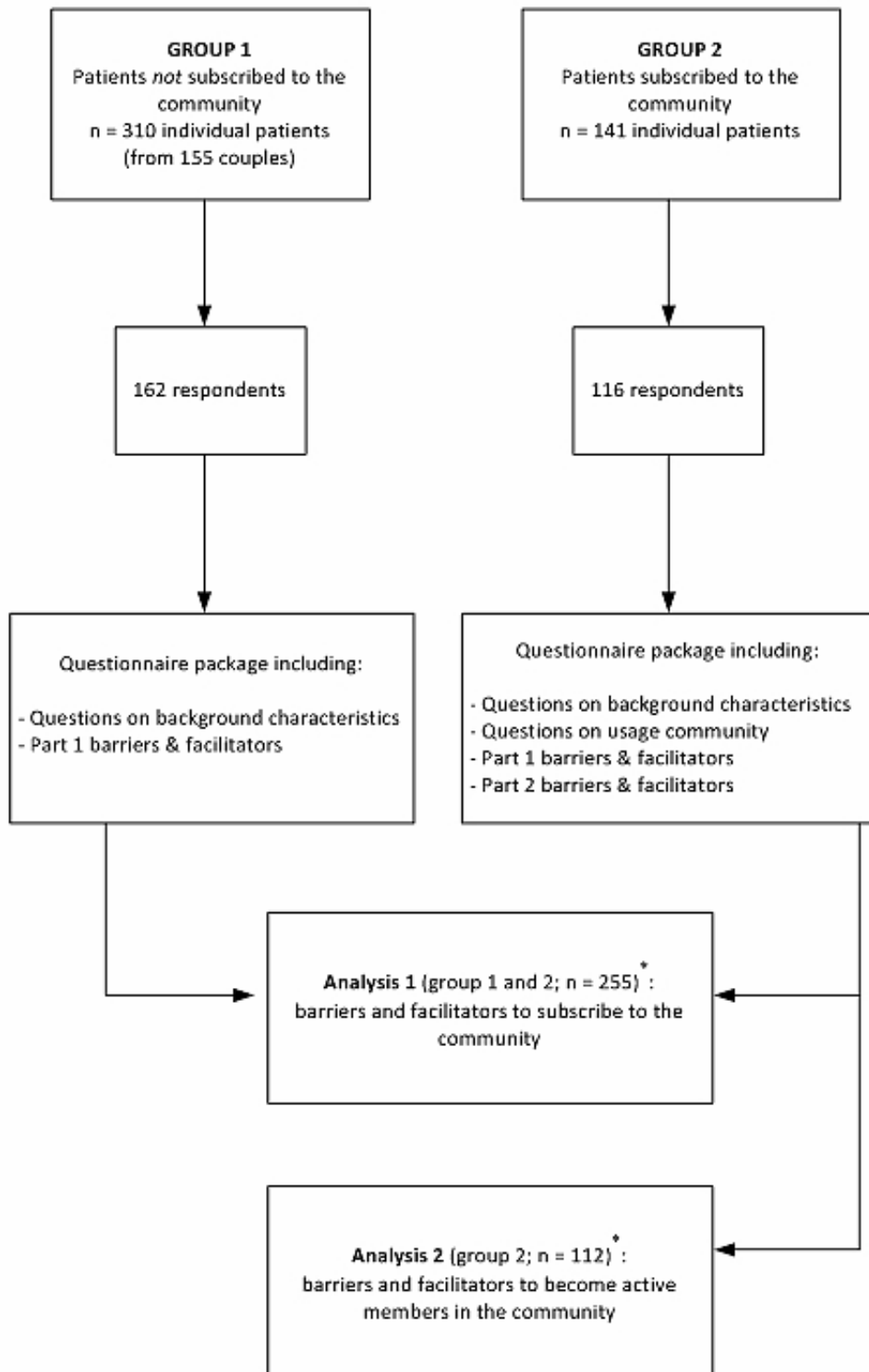


Table 1. Barriers and facilitators into domains^a resulting from the qualitative analysis.

Barriers and facilitators	Subscription		Active participation	
	Cronbach alpha ^b	Included in the analysis	Cronbach alpha ^b	Included in the analysis
Barriers				
Related to patient	.77		.64	
I'd rather call when I have a question about my treatment		Y		Y
I'd rather have face-to-face contact with my doctor/nurse		Y		Y
I don't need peer support		Y		Y
I don't need a website like this		Y		Y
Participating in this community does not fit my personality		Y		Y
I have enough knowledge about infertility and treatments		Y		Y
I have enough people (family and friends) to talk to about my feelings		Y		Y
I have little Internet experience		Y		Y
Related to intervention in general	.50		.46	
I didn't hear about it		Y		Y
I'm afraid that my privacy is not guaranteed at this website		Y		Y
I could not find the website and/or community easily		Y		Y
I experienced problems during log-on with my digital identity				Y
I don't know who the other patient members are				Y
Related to the intervention's content	n/a		.85	
Too little new information is posted on the website, such as blog messages				Y
The website does not provide much information (yet)				Y
The layout of the website doesn't invite to participate actively				Y
I think the website is poorly organized				Y
The website doesn't encourage posting comments or reactions				Y
I find using the website difficult/complicated				Y
The layout of the website consists of too much text				Y
I have to learn how to use the community				Y
Facilitators				
Related to the patient	.52		.54	
In my daily life I make use of social networking sites, such as LinkedIn or Facebook		Y		Y
I think it might be fun to use a community like this		Y		Y
I have few people to talk to about my fertility problems and feelings		Y		Y
I like to read about new facts (new treatments, research)				Y
I can help other patients by responding to questions or sharing experiences				Y
Related to the intervention	.75		.83	
Within the community I can share experiences with peers		Y		Y
Here I can easily ask questions to my physicians and nurses		Y		Y
The website has a safe impression because I have to log in using my digital identity		Y		Y
I can easily find information on this website		Y		Y

Barriers and facilitators	Subscription		Active participation	
	Cronbach alpha ^b	Included in the analysis	Cronbach alpha ^b	Included in the analysis
If I forgot to ask something during my appointment, I can do it here afterwards				Y
Here I can also find information that I wasn't looking for				Y
I know that the other members in the community are patients in the same hospital				Y
I can learn from the questions other people ask				Y
I can outlet my stories at this website				Y
The information provided at the website is reliable				Y
Related to the context	.69		.64	
The virtual infertility community is something new		Y		Y
My own doctor advised to me to use the virtual infertility community		Y		Y
The virtual infertility community is a valuable addition to usual care		Y		Y
Care becomes more patient-centered by offering this community to patients		Y		Y
Nowadays, everything is digital				
Related to the professional	n/a		.61	
Also my medical team participates actively within the community				Y
I like to read the opinion of my doctors about (new) research and treatments				Y
Because my doctors and nurses answer my questions online, it improves my relationship with them				Y

^aAccording the framework of Cabana et al [34].

Results

Overview

Figure 1 presents a schematic overview of the numbers of patients that were invited, responded, and were eligible for analyses. We invited 141 members from 1 of the 3 clinics' online health communities to participate in the study and 116 responded (82.3%). In addition, we invited both partners of 155 couples (310 individual patients) among the nonsubscribed population to participate with a response rate of 52.3% (162/310). The main reason for nonparticipation was "not willing to participate in research in general." In addition, 23 participants were removed from further analyses, because they filled out less than half of the questions on the questionnaire. Table 2 shows the background characteristics of our study population divided into 3 groups: the unsubscribed group of patients, the subscribers, and the active participants. From the total group of participants (N=255), 184 patients had heard about the online infertility community, and 111 had actually subscribed. Figure 2 presents the self-reported activity of the members of 1 of the online health communities (n=112; 1 missing). This number is the sum of the number of participants that we recruited from each of the online health communities that participated in this study.

Statistical Analyses

Tables 3 and 4 present means of sum scores, including standard deviations, for each subscale. No variables were excluded from the analyses based on collinearity.

Bivariate Relationships: Subscribers Versus Nonsubscribers

Table 3 displays the bivariate relationship between each subscale and subscription. All subscales were significantly associated with subscription in these analyses.

Bivariate Relationships: Active Versus Nonactive Groups

Table 4 presents the bivariate relationship between each subscale and active participation. All but 2 (ie, barriers related to the intervention in general and the intervention's content), were significantly associated with active participation.

Multivariate Relationships: Subscribers Versus Nonsubscribers

As presented in Table 5, in the multivariate logistic regression analysis, 5 variables predicted the willingness to subscribe to the online health community. for instance, the sum score of the barriers in the patient-related subscale significantly predicted the willingness of patients to subscribe. the higher the sum score, the more patients perceived this category as a barrier. Patients' characteristics, such as ethnicity, educational level, and average hours of Internet use per week, and context-related and

patient-related facilitators did not survive the multivariate regression analysis. the estimation of the explained variance of this multivariate regression model ($R^2=0.48$).

Multivariate Relationships: Actives Versus Nonactives

As can be seen in Table 6, 3 variables were determinants for the willingness of patients to participate actively within the

online health community after subscription. for example, the sum score of intervention-related facilitators was associated significantly with active participation within the online infertility community. Other patients' characteristics did not survive the multivariate regression analysis ($R^2=0.39$).

Table 2. Participants' background characteristics divided in three groups (unsubscribed, subscribed, and participation groups).

Demographic and treatment characteristics	Unsubscribed (n=134)	Subscribed (n=121)	Active (n=74)
Gender, n (%)			
Male	54 (40.6)	12 (9.8)	3 (4.4)
Female	80 (59.4)	109 (90.2)	71 (95.6)
Age (years), mean (SD)	33.3 (6.1)	33.4 (5.4)	32.2 (3.8)
Ethnic background,^a n (%)			
Dutch	124 (93.0)	113 (93.4)	70 (94.1)
Non-Dutch	10 (7.0)	8 (6.6)	4 (5.9)
Level of education,^b n (%)			
Low-middle	62 (46.2)	43 (35.8)	30 (41.2)
High	72 (53.8)	78 (64.2)	44 (58.8)
Duration of infertility (years), mean (SD)	2.9 (1.9)	3.4 (2.3)	3.8 (2.7)
Diagnosis, n (%)			
Male factor ^c	43 (32.2)	43 (35.7)	27 (36.8)
Female factor ^d	38 (28.7)	33 (27.7)	21 (27.9)
Both ^e	19 (14.0)	15 (12.5)	7 (8.8)
Unexplained	34 (25.2)	27 (22.3)	11 (14.7)
Treatment type, n (%)			
No treatment yet	25 (18.6)	7 (6.0)	2 (3.0)
ART ^f	58 (43.3)	85 (70.2)	60 (81.0)
non-ART ^g	50 (37.1)	29 (23.8)	12 (16.0)
Characteristics related to Internet use			
Internet use per week (hours), mean (SD)	17.1 (13.7)	18.9 (13.4)	19.3 (14.1)
Appreciation community (1-10), mean (SD)	8.2 (1.2)	8.7 (1.0)	9.0 (1.0)

^aFor ethnic background we used the Statistics Bureau Netherlands classification. This Dutch governmental institution classifies ethnicity according to citizens' country of birth and to that of their parents. Immigrants include both those who are foreign-born (first generation) and those who have at least 1 foreign-born parent (second generation). Categories were: (1) native Dutch, (2) Western or westernized origin (Europe, the United States, Canada, Australia, New Zealand, Japan, and Israel), (3) non-Western origin, immigrants from remaining countries, including Morocco, Surinam, and Turkey.

^bLow-middle: primary or lower vocational education and secondary or intermediate vocational education; high: higher professional education or university.

^cLow semen quality.

^dIrregular ovulation, polycystic ovary syndrome, tubal factor, severe endometriosis, mucus hostility.

^eBoth male and female infertility diagnosis found.

^fAssisted reproductive technology (ART) encompassed IVF, ICSI, cryopreservation, and testicular sperm extraction.

^gNon-ART included ovulation induction and intrauterine insemination with or without controlled ovarian stimulation.

Figure 2. Types of users by self-reported activity according to a participation ladder.

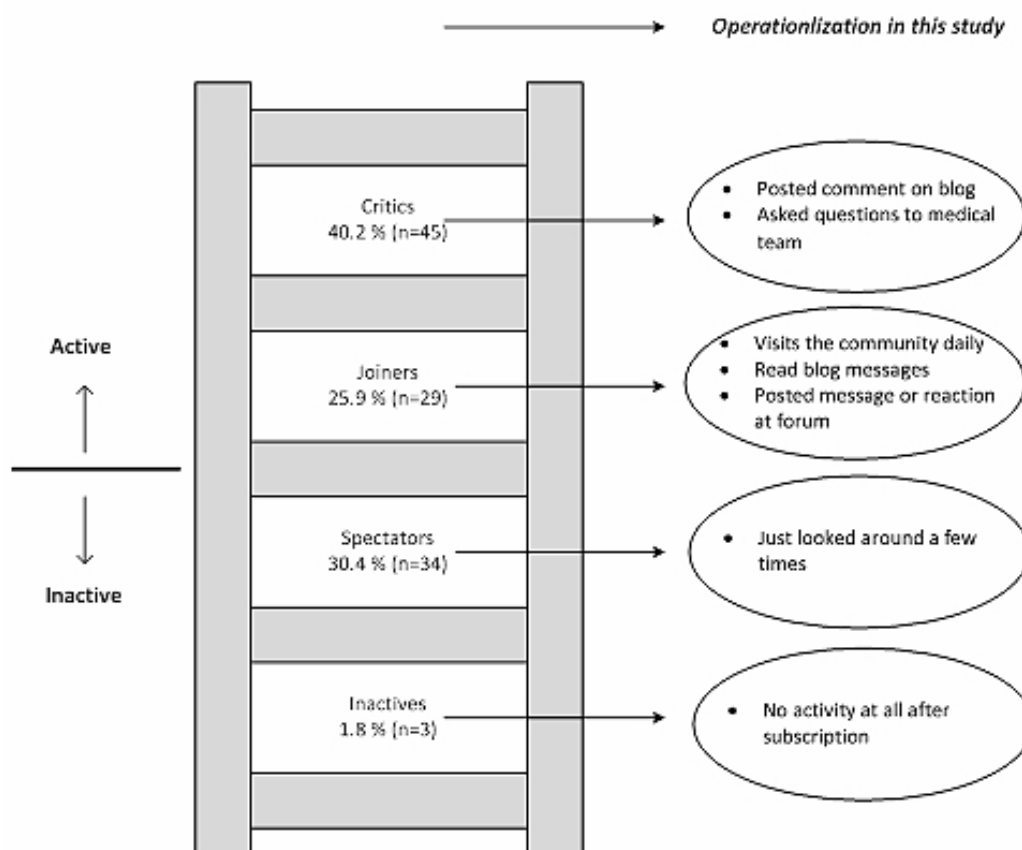


Table 3. Means (range 1-4), standard deviations, and bivariate relationships of subscribers versus nonsubscribers.

Subscales	Mean sum score ^a (SD)		Bivariate relationship		
	Subscribers (n=121)	Nonsubscribers (n=134)	OR	95% CI	P value
Barriers					
Related to the patient	1.71 (0.43)	1.98 (0.63)	0.40	0.25-0.65	<.001
Related to the intervention in general	1.41 (0.53)	1.81 (0.76)	0.39	0.26-0.59	<.001
Related to the intervention's content ^b	N/A	N/A	N/A		
Facilitators					
Related to the patient	2.19 (0.71)	1.95 (0.74)	1.59	1.12-2.27	.009
Related to the intervention	2.97 (0.67)	2.60 (0.84)	2.31	1.81-4.85	.001
Related to the context	2.67 (0.58)	2.40 (0.70)	1.89	1.27-2.83	.002
Related to the professional ^b	N/A	N/A	N/A		

^aMean sum score calculated as the mean score of each individual item divided by the number of items within the category.

^bN/A: this subscale was only used in analysis of active participation.

Table 4. Means (range 1-4), standard deviations, and bivariate relationships of nonactive versus active users.

Subscales	Mean sum score ^a (SD)		Bivariate relationship		
	Nonactives (n=37)	Actives (n=74)	OR	95% CI	P value
Barriers					
Related to the patient	1.92 (0.52)	1.57 (0.52)	0.22	0.08-0.57	.002
Related to the intervention in general	1.56 (0.57)	1.42 (0.42)	0.47	0.20-1.07	.07
Related to the intervention's content	1.71 (0.59)	1.63 (0.59)	0.63	0.33-1.22	.17
Facilitators					
Related to the patient	2.09 (0.65)	2.56 (0.55)	3.12	1.57-6.21	<.001
Related to the intervention	2.56 (0.67)	3.07 (0.52)	5.32	2.43-11.67	<.001
Related to the context	2.45 (0.53)	2.81 (0.59)	2.61	1.30-5.26	.007
Related to the professional	2.51 (0.72)	2.91 (0.68)	2.60	1.42-4.77	.002

^aMean sum score calculated as the mean score of each individual item divided by the number of items within the category.

Table 5. Multivariate relationship of background characteristics and sum scores of barriers and facilitators to subscribe to the online health community.

Independent variable	OR	95% CI	P value	Interpretation
Female	10.52	1.55-71.41	.02	Women more likely to subscribe than men.
IVF treatment	3.18	1.28-7.94	.01	IVF-treated patients more likely to subscribe than non-IVF-treated patients.
Duration of infertility (years)	1.35	1.09-1.69	.007	The longer the patient's wish for a child, the more likely they will subscribe.
Patient-related barriers	0.20	0.08-0.54	<.001	Patients perceiving patient-related barriers (eg, rather face-to-face) are less willing to subscribe.
Intervention-related facilitators	2.45	1.14-5.27	.02	Patients perceiving intervention-related facilitators are more likely they are to subscribe.

Table 6. Multivariate relationship of background characteristics and sum scores of barriers and facilitators to participate actively within the online health community after subscription.

Independent variable	OR	95% CI	P value	Interpretation
Age	0.86	0.76-0.97	.02	The younger the patients, the more likely that they will participate.
Duration of infertility (years)	1.48	1.09-2.02	.01	The longer the patient's wish for a child, the more likely they will participate.
Intervention-related facilitators	5.79	2.40-13.98	<.001	Patients perceiving intervention-related facilitators are more likely they are to participate actively.

Discussion

Principal Findings

In this study, we identified barriers and facilitators for subscription and for active participation in an online health community offered in addition to usual fertility care. Subscription appeared to be associated with several patients' background characteristics, patient-related barriers, and intervention-related facilitators. After subscription, determinants for active participation consisted of participant's age, length of infertility, and aspects related to characteristics of the online health community itself. To the best of our knowledge, this study is unique because we analyzed the barriers and facilitators for

using an Internet intervention into different phases. This provided more detailed information for future implementation strategies, which should take into account these different phases [31].

Meaning of the Study

This study provides directions on developing a targeted strategy to engage patients, in terms of subscription and active participation, in the online health community as part of the implementation of an online health community [33].

We found that intervention-related characteristics, such as sharing experiences and finding relevant information, facilitated patients' decisions to subscribe to the online health community and, thus, appealed to most of their needs. However, this did

not account for all patients. Our results also show that patient-related barriers are strongly associated with subscription: the more patient-related barriers a patient perceives, the less likely it is that he or she will subscribe. This category consists mostly of internal motivational barriers (eg, no added value) instead of external motivational barriers (eg, lacking correct skills) [36]. It could be the case that a number of people do not feel a fit with their personality. It then could be argued whether we should put too much effort into engaging people who cannot be motivated. However, an implicit explanation of our finding could be based on underlying high anxiety levels, which is not uncommon among infertile patients [37]. Anxious patients generally focus on completing simple tasks of daily living and possibly may not believe that they would benefit from an Internet intervention that comes on top of everything else [38]. However, these patients often have more need for reliable information and support from staff and peers [9], which can be provided by the online community. Therefore, we might need to spend more time identifying patients who might benefit and promoting the community actively among them. In addition, we should evaluate their experiences to optimize the community's content.

Furthermore, our results show that these patients were primarily female, undergoing IVF treatment, or had a longer duration of childlessness. Based on these results, it may seem clear-cut that we should focus on these groups of patients, but because of the cross-sectional design of our study, it is unknown in what way we should interpret the direction of this association. Either patients meeting these characteristics have more need for an online health community than, for instance, men or patients undergoing non-IVF treatments, or the way in which the content of the online health community is promoted only appeals to this subgroup. For instance, there are gender differences in needs, the experience of infertility, and strategies for coping with fertility-related problems, although infertility is considered a couples' condition [12,35,39]. Men tend to adopt task-oriented interaction styles [40] and consequently place greater importance on (medical) information than on emotional support groups in contrast to women [41-43]. Furthermore, it is known that patients undergoing diagnostic assessments or a first IUI treatment cycle also have great information needs [44] and suffer from the same emotional impact of being infertile as IVF patients [7,44,45]. Therefore, our results might reflect a lack of acknowledgment of the burden of treatment for men and non-IVF patients, which is still present in infertility services. Thus, the online infertility community could have been unintentionally promoted more prominently among IVF-treated and female patients. In our study, 24% of patients had not heard about the community. This might jeopardize equitability of care, which is also an important component of present-day high-quality care. The Internet has the capability to reach many people at the same time. However, clinics should assess the needs and expectations of different specified target groups within their patient population to tailor the promotion strategy of the online health community more appropriately to these groups. We would generally expect that the process of tailoring would make more content relevant to more people. Clinics should make sure they do not rule out certain subgroups, such

as men, in their strategy to promote the community, especially in terms of equitability of care.

In this study, we also investigated those factors that could contribute to active participation within the online health community after subscription. We know from many studies that attrition afterwards is often very high [22,28]. Previous studies have shown that Internet-based interventions only have a fair chance to be effective if members are active participants [29,30]. In our study, almost 70% of subscribers participated actively, which is a fairly high amount. Age and length of infertility were associated with active participation, although these were not strong predictors (given their 95% CIs approaching 1.00). Furthermore, echoing other studies' results, this study found that intervention-related characteristics play an important role in facilitating active participation in 2 ways. First, the types of technologies used in the community, such as blogs, forums, and wikis, make up the interactive element of the intervention through which patients can share experiences with others and communicate with their doctors. These types of technologies are believed to increase participation and reduce attrition because people get a greater feeling of engagement to the online health community [18,30,46-48]. This is confirmed in our study. Second, the content of the community—a combination of peer-to-peer communication, patient-to-professional communication, and information provision—facilitated active participation, which implies that it fulfilled subscribers' needs generating value for them. This underlines that it is important to tailor the intervention to patient's needs.

Although the subscale professional-related facilitators, including active participation from the medical team in the online community, did not survive the multivariate regression analysis, it appeared to have a fairly strong bivariate relationship to active participation. This is in-line with findings in some previous studies: frequent news updates and active participation from clinicians attract patients [47-50]. However, clinicians do also perceive barriers for participating within these types of Internet-based interventions [49,51-54], such as time constraints or lack of knowledge of benefits. Future studies should investigate what specific barriers and facilitators clinicians experience as a next step in the development of a tailored implementation strategy.

Limitations and Strengths

A strength of our study is that the questionnaire was based on the factors identified by qualitative research. This method assures that the survey is not testing the authors' personal hypothesis, but represents the complete spectrum of the factors related to adoption of an online infertility community. Another strong point is the fact that we obtained a representative sample of participants and questioned them in a real-life setting instead of an experimental one. The online health community was added to usual care in the clinic they visited. This contributes to the validity of our findings. A difficulty of this study relates to the question whether it can be generalized to other contexts, such as other clinics or other countries. Another context might bring about other barriers and facilitators for the adoption of this intervention. Nevertheless, most factors can be considered universal and probably not specifically related to the Dutch care

setting, a second limitation is that we were not able to measure patients' activity within the online health community objectively, but used self-reported activity instead. Third, it would have been interesting to include every single item from the questionnaire into the regression model. However, our sample size was too small because we needed at least 20 patients for each additional independent variable in the model [55]. Therefore, we narrowed the number of independent variables by using subscales based on rigorously performed qualitative analysis.

Conclusions

In this questionnaire study, we searched for factors that are associated with subscription to and subsequent active participation in an online fertility community in addition to usual care delivery. We concluded that being female, undergoing IVF treatment, patient-related barriers, and intervention-related facilitators are associated with subscription to the community. Participant's age, length of infertility, and intervention-related characteristics facilitated the active participation of these subscribers within the online community. These results imply that involving patients and their needs into the promotion strategy, the community's design, and the implementation plan are crucial.

Acknowledgments

The authors thank Mr J Koetsenruijter for his support in the statistical analyses. This study was supported by the Department of Obstetrics and Gynecology of the Radboud University Nijmegen Medical Center (Nijmegen, the Netherlands). The intervention under investigation was financially supported by the Dutch Ministry of Health, Welfare and Sport (the Hague, the Netherlands).

Authors' Contributions

JA designed the study, developed and managed the main database, interpreted the analysis, and drafted the paper. AdB, BC, and PvL contributed to interpretation of data and revised critically for important intellectual content. AdB also contributed to data analysis and management. MF, WN, and JK designed the study, interpreted the data, and drafted the paper. All authors gave final approval of the version to be published.

Conflicts of Interest

None declared.

References

1. Bensing J. Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine. *Patient Educ Couns* 2000 Jan;39(1):17-25. [Medline: [11013544](#)]
2. Berwick DM. What 'patient-centered' should mean: confessions of an extremist. *Health Aff (Millwood)* 2009;28(4):w555-w565 [FREE Full text] [doi: [10.1377/hlthaff.28.4.w555](#)] [Medline: [19454528](#)]
3. Coulter A. Patients' views of the good doctor. *BMJ* 2002 Sep 28;325(7366):668-669 [FREE Full text] [Medline: [12351342](#)]
4. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. *BMJ* 2003 Apr 19;326(7394):877-879 [FREE Full text] [doi: [10.1136/bmj.326.7394.877](#)] [Medline: [12702627](#)]
5. Malin M, Hemmink E, Räikkönen O, Sihvo S, Perälä ML. What do women want? Women's experiences of infertility treatment. *Soc Sci Med* 2001 Jul;53(1):123-133. [Medline: [11380158](#)]
6. Dancet EA, Nelen WL, Sermeus W, De Leeuw L, Kremer JA, D'Hooghe TM. The patients' perspective on fertility care: a systematic review. *Hum Reprod Update* 2010;16(5):467-487 [FREE Full text] [doi: [10.1093/humupd/dmq004](#)] [Medline: [20223789](#)]
7. Schmidt L, Holstein BE, Boivin J, Sångren H, Tjørnhøj-Thomsen T, Blaabjerg J, et al. Patients' attitudes to medical and psychosocial aspects of care in fertility clinics: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) Research Programme. *Hum Reprod* 2003 Mar;18(3):628-637 [FREE Full text] [Medline: [12615837](#)]
8. Redshaw M, Hockley C, Davidson LL. A qualitative study of the experience of treatment for infertility among women who successfully became pregnant. *Hum Reprod* 2007 Jan;22(1):295-304 [FREE Full text] [doi: [10.1093/humrep/del344](#)] [Medline: [16959809](#)]
9. Mourad SM, Hermens RP, Cox-Witbraad T, Grol RP, Nelen WL, Kremer JA. Information provision in fertility care: a call for improvement. *Hum Reprod* 2009 Jun;24(6):1420-1426 [FREE Full text] [doi: [10.1093/humrep/dep029](#)] [Medline: [19224886](#)]
10. van Peperstraten A, Nelen W, Grol R, Zielhuis G, Adang E, Stalmeier P, et al. The effect of a multifaceted empowerment strategy on decision making about the number of embryos transferred in in vitro fertilisation: randomised controlled trial. *BMJ* 2010;341:c2501 [FREE Full text] [Medline: [20884700](#)]
11. van Empel IW, Aarts JW, Cohlen BJ, Huppelschoten DA, Laven JS, Nelen WL, et al. Measuring patient-centredness, the neglected outcome in fertility care: a random multicentre validation study. *Hum Reprod* 2010 Oct;25(10):2516-2526 [FREE Full text] [doi: [10.1093/humrep/deq219](#)] [Medline: [20719811](#)]

12. Haagen EC, Tuil W, Hendriks J, de Bruijn RP, Braat DD, Kremer JA. Current Internet use and preferences of IVF and ICSI patients. *Hum Reprod* 2003 Oct;18(10):2073-2078 [[FREE Full text](#)] [Medline: [14507823](#)]
13. Weissman A, Gotlieb L, Ward S, Greenblatt E, Casper RF. Use of the internet by infertile couples. *Fertil Steril* 2000 Jun;73(6):1179-1182. [Medline: [10856479](#)]
14. Dancet EA, Van Empel IW, Rober P, Nelen WL, Kremer JA, D'Hooghe TM. Patient-centred infertility care: a qualitative study to listen to the patient's voice. *Hum Reprod* 2011 Apr;26(4):827-833 [[FREE Full text](#)] [doi: [10.1093/humrep/der022](#)] [Medline: [21317152](#)]
15. 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](#)] [Medline: [18725354](#)]
16. Shirky C. *Here Comes Everybody: How Change Happens When People Come Together*. London: Penguin Group(CA); 2008.
17. Whittington K, Cook J, Barratt C, Jenkins J. Can the Internet widen participation in reproductive medicine education for professionals? *Hum Reprod* 2004 Aug;19(8):1800-1805 [[FREE Full text](#)] [doi: [10.1093/humrep/deh333](#)] [Medline: [15178660](#)]
18. Aarts JW, van den Haak P, Nelen WL, Tuil WS, Faber MJ, Kremer JA. Patient-focused internet interventions in reproductive medicine: a scoping review. *Hum Reprod Update* 2012;18(2):211-227 [[FREE Full text](#)] [doi: [10.1093/humupd/dmr045](#)] [Medline: [22108381](#)]
19. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. *Qual Health Res* 2008 Mar;18(3):405-417. [doi: [10.1177/1049732307313429](#)] [Medline: [18235163](#)]
20. Baker TB, Gustafson DH, Shaw B, Hawkins R, Pingree S, Roberts L, et al. Relevance of CONSORT reporting criteria for research on eHealth interventions. *Patient Educ Couns* 2010 Dec;81 Suppl:S77-S86 [[FREE Full text](#)] [doi: [10.1016/j.pec.2010.07.040](#)] [Medline: [20843621](#)]
21. Demiris G. The diffusion of virtual communities in health care: concepts and challenges. *Patient Educ Couns* 2006 Aug;62(2):178-188. [doi: [10.1016/j.pec.2005.10.003](#)] [Medline: [16406472](#)]
22. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 15;328(7449):1166 [[FREE Full text](#)] [doi: [10.1136/bmj.328.7449.1166](#)] [Medline: [15142921](#)]
23. Tuil WS, Verhaak CM, Braat DD, de Vries Robbé PF, Kremer JA. Empowering patients undergoing in vitro fertilization by providing Internet access to medical data. *Fertil Steril* 2007 Aug;88(2):361-368. [doi: [10.1016/j.fertnstert.2006.11.197](#)] [Medline: [17416366](#)]
24. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. *Patient Educ Couns* 2009 Jan;74(1):61-69. [doi: [10.1016/j.pec.2008.07.044](#)] [Medline: [18778909](#)]
25. Malik SH, Coulson NS. Computer-mediated infertility support groups: an exploratory study of online experiences. *Patient Educ Couns* 2008 Oct;73(1):105-113. [doi: [10.1016/j.pec.2008.05.024](#)] [Medline: [18639409](#)]
26. Nambisan P. Information seeking and social support in online health communities: impact on patients' perceived empathy. *J Am Med Inform Assoc* 2011 May 1;18(3):298-304 [[FREE Full text](#)] [doi: [10.1136/amiainl-2010-000058](#)] [Medline: [21486888](#)]
27. Murray E, Burns J, May C, Finch T, O'Donnell C, Wallace P, et al. Why is it difficult to implement e-health initiatives? A qualitative study. *Implement Sci* 2011;6:6 [[FREE Full text](#)] [doi: [10.1186/1748-5908-6-6](#)] [Medline: [21244714](#)]
28. Nijland N, van Gemert-Pijnen JE, Boer H, Stehouder MF, Seydel ER. Increasing the use of e-consultation in primary care: results of an online survey among non-users of e-consultation. *Int J Med Inform* 2009 Oct;78(10):688-703. [doi: [10.1016/j.ijmedinf.2009.06.002](#)] [Medline: [19625210](#)]
29. Van 't Riet J, Crutzen R, De Vries H. Investigating predictors of visiting, using, and revisiting an online health-communication program: a longitudinal study. *J Med Internet Res* 2010;12(3):e37 [[FREE Full text](#)] [doi: [10.2196/jmir.1345](#)] [Medline: [20813716](#)]
30. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. *J Med Internet Res* 2010;12(4):e71 [[FREE Full text](#)] [doi: [10.2196/jmir.1338](#)] [Medline: [21169160](#)]
31. Chiu TM, Eysenbach G. Stages of use: consideration, initiation, utilization, and outcomes of an internet-mediated intervention. *BMC Med Inform Decis Mak* 2010;10:73 [[FREE Full text](#)] [doi: [10.1186/1472-6947-10-73](#)] [Medline: [21092275](#)]
32. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. *Med J Aust* 2004 Mar 15;180(6 Suppl):S57-S60. [Medline: [15012583](#)]
33. MijnZorgnet [MyCareNet]. URL: <https://www.mijnzorgnet.nl/welcome.aspx> [accessed 2011-12-22] [[WebCite Cache ID 647Vcyksz](#)]
34. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999 Oct 20;282(15):1458-1465. [Medline: [10535437](#)]
35. Li C. Empowered. 2007 Apr 23. Forrester's new Social Technographics report URL: http://forrester.typepad.com/groundswell/2007/04/forresters_new_.html [accessed 2013-07-16] [[WebCite Cache ID 6I9Wqsey](#)]

36. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. *J Med Internet Res* 2011;13(3):e71 [FREE Full text] [doi: [10.2196/jmir.1603](https://doi.org/10.2196/jmir.1603)] [Medline: [21959968](https://pubmed.ncbi.nlm.nih.gov/21959968/)]
37. Verhaak CM, Lintsen AM, Evers AW, Braat DD. Who is at risk of emotional problems and how do you know? Screening of women going for IVF treatment. *Hum Reprod* 2010 May;25(5):1234-1240 [FREE Full text] [doi: [10.1093/humrep/deq054](https://doi.org/10.1093/humrep/deq054)] [Medline: [20228392](https://pubmed.ncbi.nlm.nih.gov/20228392/)]
38. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. *J Med Internet Res* 2012;14(3):e91 [FREE Full text] [doi: [10.2196/jmir.2100](https://doi.org/10.2196/jmir.2100)] [Medline: [22743581](https://pubmed.ncbi.nlm.nih.gov/22743581/)]
39. Mo PK, Malik SH, Coulson NS. Gender differences in computer-mediated communication: a systematic literature review of online health-related support groups. *Patient Educ Couns* 2009 Apr;75(1):16-24. [doi: [10.1016/j.pec.2008.08.029](https://doi.org/10.1016/j.pec.2008.08.029)] [Medline: [19233586](https://pubmed.ncbi.nlm.nih.gov/19233586/)]
40. Pasch LA, Dunkel-Schetter C, Christensen A. Differences between husbands' and wives' approach to infertility affect marital communication and adjustment. *Fertil Steril* 2002 Jun;77(6):1241-1247. [Medline: [12057735](https://pubmed.ncbi.nlm.nih.gov/12057735/)]
41. Malik S, Coulson N. Computer-mediated infertility support groups: an exploratory study of online experiences. *Patient Educ Couns* 2008 Oct;73(1):105-113. [doi: [10.1016/j.pec.2008.05.024](https://doi.org/10.1016/j.pec.2008.05.024)] [Medline: [18639409](https://pubmed.ncbi.nlm.nih.gov/18639409/)]
42. Kiss A, Meryn S. Effect of sex and gender on psychosocial aspects of prostate and breast cancer. *BMJ* 2001 Nov 3;323(7320):1055-1058 [FREE Full text] [Medline: [11691767](https://pubmed.ncbi.nlm.nih.gov/11691767/)]
43. Owen JE, Klapow JC, Roth DL, Tucker DC. Use of the internet for information and support: disclosure among persons with breast and prostate cancer. *J Behav Med* 2004 Oct;27(5):491-505. [Medline: [15675637](https://pubmed.ncbi.nlm.nih.gov/15675637/)]
44. Ezer H, Chachamovich JL, Chachamovich E. Do men and their wives see it the same way? Congruence within couples during the first year of prostate cancer. *Psychooncology* 2011 Feb;20(2):155-164. [doi: [10.1002/pon.1724](https://doi.org/10.1002/pon.1724)] [Medline: [20878828](https://pubmed.ncbi.nlm.nih.gov/20878828/)]
45. Haagen EC, Hermens RP, Nelen WL, Braat DD, Kremer JA, Grol RP. Subfertile couples' negative experiences with intrauterine insemination care. *Fertil Steril* 2008 Apr;89(4):809-816. [doi: [10.1016/j.fertnstert.2007.04.005](https://doi.org/10.1016/j.fertnstert.2007.04.005)] [Medline: [17706205](https://pubmed.ncbi.nlm.nih.gov/17706205/)]
46. Aarts JW, van Empel IW, Boivin J, Nelen WL, Kremer JA, Verhaak CM. Relationship between quality of life and distress in infertility: a validation study of the Dutch FertiQoL. *Hum Reprod* 2011 May;26(5):1112-1118 [FREE Full text] [doi: [10.1093/humrep/der051](https://doi.org/10.1093/humrep/der051)] [Medline: [21372046](https://pubmed.ncbi.nlm.nih.gov/21372046/)]
47. Crutzen R, Cyr D, de Vries NK. Bringing loyalty to e-Health: theory validation using three internet-delivered interventions. *J Med Internet Res* 2011;13(3):e73 [FREE Full text] [doi: [10.2196/jmir.1837](https://doi.org/10.2196/jmir.1837)] [Medline: [21946128](https://pubmed.ncbi.nlm.nih.gov/21946128/)]
48. Wasson JH, Forsberg HH, Lindblad S, Mazowita G, McQuillen K, Nelson EC. The medium is the (health) measure: patient engagement using personal technologies. *J Ambul Care Manage* 2012;35(2):109-117. [doi: [10.1097/JAC.0b013e31824a235e](https://doi.org/10.1097/JAC.0b013e31824a235e)] [Medline: [22415284](https://pubmed.ncbi.nlm.nih.gov/22415284/)]
49. Resnick PJ, Janney AW, Buis LR, Richardson CR. Adding an online community to an internet-mediated walking program. Part 2: strategies for encouraging community participation. *J Med Internet Res* 2010;12(4):e72 [FREE Full text] [doi: [10.2196/jmir.1339](https://doi.org/10.2196/jmir.1339)] [Medline: [21169161](https://pubmed.ncbi.nlm.nih.gov/21169161/)]
50. Nordqvist C, Hanberger L, Timpka T, Nordfeldt S. Health professionals' attitudes towards using a Web 2.0 portal for child and adolescent diabetes care: qualitative study. *J Med Internet Res* 2009;11(2):e12 [FREE Full text] [doi: [10.2196/jmir.1152](https://doi.org/10.2196/jmir.1152)] [Medline: [19403464](https://pubmed.ncbi.nlm.nih.gov/19403464/)]
51. Nordfeldt S, Hanberger L, Berterö C. Patient and parent views on a Web 2.0 Diabetes Portal--the management tool, the generator, and the gatekeeper: qualitative study. *J Med Internet Res* 2010;12(2):e17 [FREE Full text] [doi: [10.2196/jmir.1267](https://doi.org/10.2196/jmir.1267)] [Medline: [20511179](https://pubmed.ncbi.nlm.nih.gov/20511179/)]
52. Koivunen M, Hätönen H, Välimäki M. Barriers and facilitators influencing the implementation of an interactive Internet-portal application for patient education in psychiatric hospitals. *Patient Educ Couns* 2008 Mar;70(3):412-419. [doi: [10.1016/j.pec.2007.11.002](https://doi.org/10.1016/j.pec.2007.11.002)] [Medline: [18079085](https://pubmed.ncbi.nlm.nih.gov/18079085/)]
53. Kerr C, Murray E, Noble L, Morris R, Bottomley C, Stevenson F, et al. The potential of Web-based interventions for heart disease self-management: a mixed methods investigation. *J Med Internet Res* 2010;12(4):e56 [FREE Full text] [doi: [10.2196/jmir.1438](https://doi.org/10.2196/jmir.1438)] [Medline: [21156471](https://pubmed.ncbi.nlm.nih.gov/21156471/)]
54. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. *Med Care Res Rev* 2007 Dec;64(6):650-672. [doi: [10.1177/1077558707305942](https://doi.org/10.1177/1077558707305942)] [Medline: [17717378](https://pubmed.ncbi.nlm.nih.gov/17717378/)]
55. Field A. *Discovering Statistics Using SPSS*. London: Sage Publications; 2005.

Abbreviations

- ART:** assisted reproductive technology
- ICSI:** intracytoplasmic sperm injection
- IUI:** intrauterine insemination
- IVF:** in vitro fertilization
- OI:** ovulation induction

OR: odds ratio

Edited by G Eysenbach; submitted 01.03.12; peer-reviewed by T Chiu; comments to author 09.07.12; revised version received 21.10.12; accepted 07.04.13; published 25.08.13.

Please cite as:

*Aarts JWM, Faber MJ, den Boogert AG, Cohlen BJ, van der Linden PJQ, Kremer JAM, Nelen WLDM
Barriers and Facilitators for the Implementation of an Online Clinical Health Community in Addition to Usual Fertility Care: A
Cross-Sectional Study
J Med Internet Res 2013;15(8):e163
URL: <http://www.jmir.org/2013/8/e163/>
doi: [10.2196/jmir.2098](https://doi.org/10.2196/jmir.2098)
PMID: [23996964](https://pubmed.ncbi.nlm.nih.gov/23996964/)*

©Johanna WM Aarts, Marjan J Faber, Anne G den Boogert, Ben J Cohlen, Paul JQ van der Linden, Jan AM Kremer, Willianne LDM Nelen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 25.08.2013. 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 Determinants of Consumer Mobile Health Usage Intentions, Assimilation, and Channel Preferences

Arun Rai¹, MS, MBA, PhD; Liwei Chen², MS; Jessica Pye², MA; Aaron Baird³, MBA, PhD

¹Center for Process Innovation and Department of Computer Information Systems, J Mack Robinson College of Business, Georgia State University, Atlanta, GA, United States

²Center for Process Innovation, J Mack Robinson College of Business, Georgia State University, Atlanta, GA, United States

³Center for Health Information Technology and Institute of Health Administration, J Mack Robinson College of Business, Georgia State University, Atlanta, GA, United States

Corresponding Author:

Arun Rai, MS, MBA, PhD

Center for Process Innovation and Department of Computer Information Systems

J Mack Robinson College of Business

Georgia State University

PO Box 4015

Atlanta, GA, 30302-4015

United States

Phone: 1 4044137857

Fax: 1 4044137867

Email: arunrai@gsu.edu

Abstract

Background: Consumer use of mobile devices as health service delivery aids (mHealth) is growing, especially as smartphones become ubiquitous. However, questions remain as to how consumer traits, health perceptions, situational characteristics, and demographics may affect consumer mHealth usage intentions, assimilation, and channel preferences.

Objective: We examine how consumers' personal innovativeness toward mobile services (PIMS), perceived health conditions, health care availability, health care utilization, demographics, and socioeconomic status affect their (1) mHealth usage intentions and extent of mHealth assimilation, and (2) preference for mHealth as a complement or substitute for in-person doctor visits.

Methods: Leveraging constructs from research in technology acceptance, technology assimilation, consumer behavior, and health informatics, we developed a cross-sectional online survey to study determinants of consumers' mHealth usage intentions, assimilation, and channel preferences. Data were collected from 1132 nationally representative US consumers and analyzed by using moderated multivariate regressions and ANOVA.

Results: The results indicate that (1) 430 of 1132 consumers in our sample (37.99%) have started using mHealth, (2) a larger quantity of consumers are favorable to using mHealth as a complement to in-person doctor visits (758/1132, 66.96%) than as a substitute (532/1132, 47.00%), and (3) consumers' PIMS and perceived health conditions have significant positive direct influences on mHealth usage intentions, assimilation, and channel preferences, and significant positive interactive influences on assimilation and channel preferences. The independent variables within the moderated regressions collectively explained 59.70% variance in mHealth usage intentions, 60.41% in mHealth assimilation, 34.29% in preference for complementary use of mHealth, and 45.30% in preference for substitutive use of mHealth. In a follow-up ANOVA examination, we found that those who were more favorable toward using mHealth as a substitute for in-person doctor visits than as a complement indicated stronger intentions to use mHealth ($F_{1,702}=20.14$, $P<.001$) and stronger assimilation of mHealth ($F_{1,702}=41.866$, $P<.001$).

Conclusions: Multiple predictors are shown to have significant associations with mHealth usage intentions, assimilation, and channel preferences. We suggest that future initiatives to promote mHealth should shift targeting of consumers from coarse demographics to nuanced considerations of individual dispositions toward mobile service innovations, complementary or substitutive channel use preferences, perceived health conditions, health services availability and utilization, demographics, and socioeconomic characteristics.

(*J Med Internet Res* 2013;15(8):e149) doi:[10.2196/jmir.2635](https://doi.org/10.2196/jmir.2635)

KEYWORDS

mobile health; consumer preferences; adoption; health information technology; multivariate analyses

Introduction

Background

Mobile health (mHealth) is defined as, “using wireless mobile communication technology to aid health services delivery” [1]. According to a recent health care market research study, 31% of US adults have used their mobile phones for accessing health information [2]. In addition, 19% of US adults who own a smartphone have at least 1 health application on their phone, with exercise, diet, and weight apps among the most popular [2]. Approximately half of the patients surveyed in a recent mHealth opinion survey believed that mHealth could increase their control over their health care, provide more convenient access to needed health information, and ultimately improve their health care costs and quality [3]. Such results are not surprising because mHealth can provide many benefits, including portable access to continuous streams of information and powerful interactive functionality driven by devices that often support a wide array of health applications [4]. However, questions remain as to what determines whether consumers will use and assimilate mHealth and whether or not channel preferences will play a significant role.

The introduction of mHealth represents a drastic shift in focus from traditional medical informatics based on industrial age concepts (eg, provider driven) to consumer health informatics based on the ubiquity of information and interconnected mobile computing infrastructure [5]. In practice, mHealth is often used for transmitting electronic medical records between medical staff and patients [6], monitoring patients remotely [6,7], sending electronic alerts for disease control [8], and providing useful applications, information, and functionality to health consumers [2]. The general category of mHealth innovations considered in this paper are typically used by consumers for activities relating to obtaining health advice (eg, the WebMD mobile app [9]), promoting compliance and adherence to medical treatments (eg, the iPharmacy Pill ID & Rx Reminder app [10]), staying connected with health care provider(s) (eg, the Mayo Clinic Patient app [11] and the eClinicalMobile app [12]), personal health management (eg, the GoMeals app [13], the Livestrong app [14], and the WellDoc app [15]), and chronic disease management (eg, the Glucose Buddy app [16] for diabetics).

Research in the mHealth context has demonstrated that intrinsic motivations facilitate mHealth adoption whereas perceived risks, such as perceived privacy risks and perceived psychological risks associated with making choices that may be regretted later, can inhibit mHealth adoption [17]. Perceptions and attitudes toward mHealth have been shown to positively affect an individual’s intention to use these types of services [18]. It has been suggested that the digital revolution brought by mobile and other technology has enriched doctor-patient communications [19]. Use of gamification in mHealth has recently been shown to increase glucose monitoring in diabetic

adolescents [20]. Studies have also examined mHealth trends and associated risks [21,22], the impact of mHealth interventions on outcomes in specific clinical areas (eg, smoking cessation [4], HIV [23], and diabetes [24]), economic implications of mHealth usage (eg, [25]), and the use of mHealth to broaden access to health care in developing countries (eg, [26,27]).

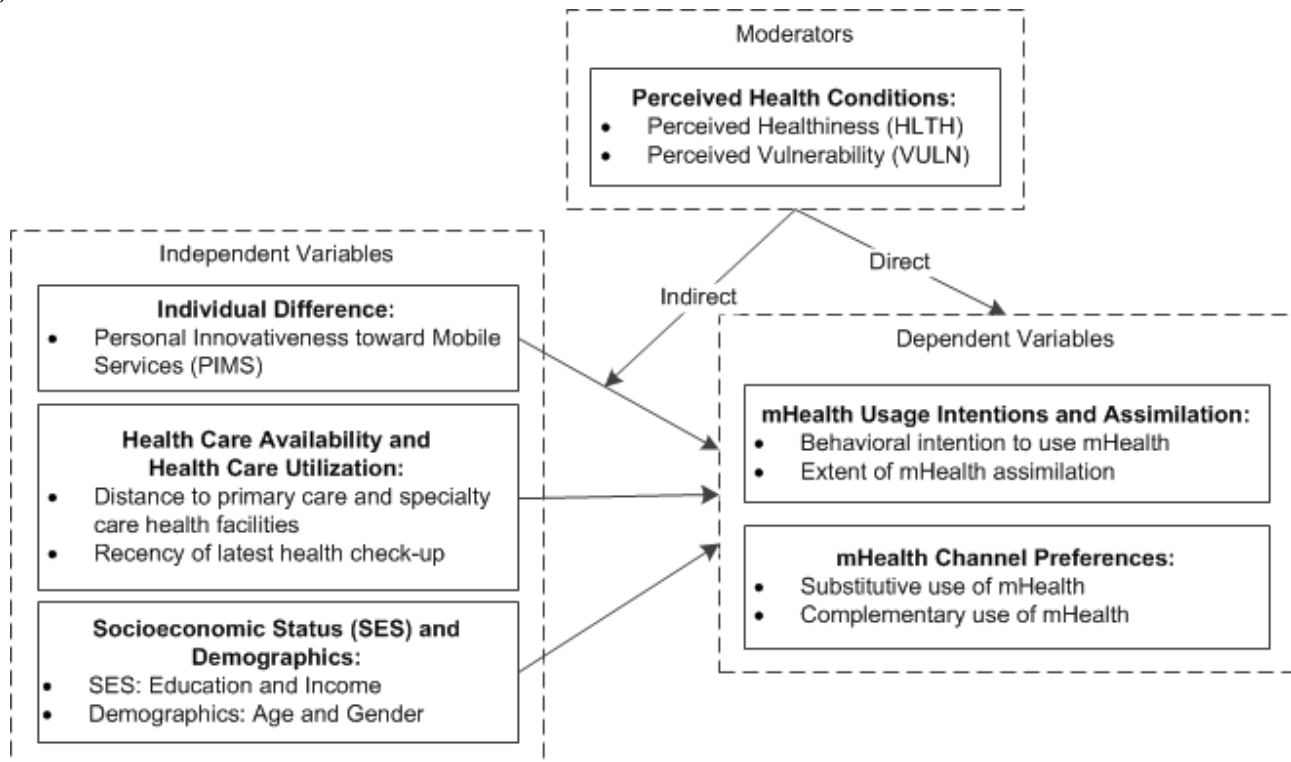
Although expectations of the transformative (and disruptive) potential of mHealth are enormous and research is expanding in this area, little is known about how this digital health care service channel is viewed by consumers, given a traditionally hands-on provider-patient direct service channel. Given recent calls for more consumer health informatics research, especially in regards to consumer information seeking needs and behaviors [28], mHealth [29], and our presently limited knowledge of how consumers’ traits and health perceptions affect consumers’ mHealth usage intentions, assimilation, and channel preferences, our study is motivated by the substantial research opportunities in this interesting and emerging space. We specifically focus on what determinants are associated with consumer mHealth usage intentions, assimilation, and channel preferences.

Theoretical Foundation and Research Model

Past work on the individual adoption of information technology (IT) has identified that consumer characteristics (eg, socioeconomic characteristics [30]), individual differences (eg, personal innovativeness [31,32]), and situational factors (eg, access to and utilization of health care services [33,34]) significantly impact IT adoption preferences. A recent systematic review of consumer health technology acceptance studies pointed out that many studies have assessed the effects of consumer traits (eg, age, income, education) on health technology acceptance, but theoretically motivated constructs, interaction effects (moderators), and health status variables have not yet been fully considered in consumer health technology acceptance studies [33]. Additionally, mHealth studies have not yet jointly examined consumer traits, health perceptions, and consumer preferences for mHealth as a substitute or complement to in-person doctor office visits.

Drawing upon technology adoption [33,35], technology assimilation [36,37], consumer behavior [38,39], and health informatics literature (eg [33,40-42]), we seek to fill this gap by focusing on determinants of consumer mHealth usage intentions, assimilation, and channel preferences in the United States. We aim to contribute to the health informatics and mHealth literature by assessing the following: (1) predictors of consumer mHealth usage intentions and assimilation, including personal innovativeness toward mobile services (PIMS), health care availability, health care utilization, socioeconomic status, and demographics, (2) consumer preferences for mHealth as a substitute or complement to in-person provider-patient interactions, and (3) the direct and interactive (moderating) effects of perceived health conditions (Figure 1).

Figure 1. Research model.



Dependent Variables: Behavioral Intention to Use mHealth and Extent of mHealth Assimilation

Based on the literature on technology acceptance [33,35] and technology assimilation [36,37], we conceptualize 2 dependent variables associated with consumers’ mHealth acceptance: (1) mHealth usage intentions (ie, intention to use mHealth for nonadopters and intention to continue using mHealth for adopters) and, (2) mHealth assimilation (ie, combined, staged measure ranging from extent of awareness to frequency of use). Self-reported behavioral intention to use an information system is a widely used dependent variable in technology acceptance research (eg, [43]) and is designed to measure intention of initial technology usage or continued usage (referred to collectively as *intention to use* for the remainder of this paper). Technology assimilation is often used as an indicator of the process of learning and applying a technology through multiple stages, ranging from very early stages of evaluation of options (awareness) to later stages of extensive use in which the technology has been incorporated into routines (frequent use) [44] (referred to collectively as *assimilation* for the remainder of this paper). The use of these 2 dependent variables in our analyses is meant to explore commonalities and differences in predictors between self-reported mHealth usage intentions and self-reported mHealth assimilation.

Dependent Variables: mHealth Substitutive and Complementary Use

Drawing from consumer behavior literature [38,39], we consider 2 dependent variables that reflect consumers’ mHealth channel preferences: (1) substitutive use (ie, the willingness to use mHealth as a replacement for in-person doctor visits) and, (2) complementary use of mHealth (ie, the willingness to use mHealth to augment in-person doctor visits). Consumer behavior

literature has suggested that offering alternative channels for service consumption (eg, self-service technologies) affords many conveniences and benefits to consumers [45,46]. However, such conveniences and benefits may come at a cost of taking the time to learn how to use the new channel, expending additional individual effort perhaps not required when interacting in-person, and concerns about overall service performance given the new production and consumption medium [47]. Given that health care has traditionally been conducted via hands-on in-person interactions, an important consideration is whether or not consumers will accept technology intermediation. Therefore, we assess consumers’ willingness to use mHealth as either a substitute or a complement to in-person doctor visits, irrespective of whether or not they are currently mHealth users.

Independent Variable: Personal Innovativeness Toward Mobile Services

We draw on marketing and information systems research to identify personal innovativeness toward mobile services (PIMS) as a key individual difference for evaluating consumer adoption of technology innovations. Based on prior work [31,48-50], we define PIMS as the degree to which an individual is willing to try out any new mobile technology service. We suggest that PIMS is a proxy for personal innovativeness toward information technology (PIIT) [31] in the context of mobile services. Prior research on PIIT has suggested that individuals who are more open to experiences with technology typically have stronger technical self-efficacy (positive beliefs and abilities associated with technical tasks, such as using a computer) [51,52]. Those with higher PIIT are often associated with higher levels technology use, such as Internet and e-commerce use [53-55]. Recent studies have found personal innovativeness to be also a positive predictor in models assessing acceptance of mobile

services [56,57]. Given that the delivery of health services on mobile platforms is currently at the very early stages of diffusion, it is likely that individuals with a predisposition for seeking out the latest innovations are more likely to initially adopt and use mHealth. Therefore, we extend the examination of PIIT to the context of mHealth using the PIMS construct.

Independent Variables: Perceived Health Conditions

Health promotion and prevention research has widely acknowledged that individual beliefs about health conditions (eg, perceived healthiness and perceived vulnerability to chronic disease) predict individuals' health behaviors (see reviews by [40,42]). We expect that consumers who feel healthier may be more open to trying health innovations. We base this expectation on evidence that those with lower perceived health may already have strong preexisting relationships with service providers (physicians and clinicians, in our case) and established therapeutic routines that may result in resistance to alternative service delivery and consumption options [46]. We also expect that people who feel more vulnerable to chronic diseases (eg, diabetes, heart disease, cancer, high blood pressure, and stroke) will have stronger needs and motivations to use health innovations. This expectation is based on evidence suggesting that those who currently report using mHealth often do so to mitigate negative long-term health consequences associated with health risk factors such as high blood pressure, obesity, inactivity, and high blood glucose levels [2].

Moderating Effects: Personal Innovativeness Toward Mobile Services and Perceived Health Conditions

Prior research on technology acceptance has shown that PIIT and compatibility with work style (CMP) typically have positive interactive (PIIT**CMP*) effects on technology usage intentions, suggesting that higher PIIT combined with higher CMP has an even greater positive impact on technology usage intentions than PIIT alone (eg, [48]). Compatibility is typically viewed as the degree of congruence between the innovation and the adopter's preferences, needs, past experience, and/or values (eg, [58,59]). Drawing upon this literature, we propose the perceived health condition of an individual as a proxy for CMP, given that perceived health conditions often reflect health care needs (eg, [60]). We expect PIMS and health conditions to have both direct and indirect (moderating) effects. Specifically, we propose that PIMS and perceived healthiness (PIMS**HLTH*) and PIMS and perceived vulnerability to chronic disease (PIMS**VULN*) will have significant interactive effects on mHealth usage intentions, assimilation, and channel preferences.

Independent Variables: Health Care Availability and Health Care Utilization

Consumers' situational characteristics have been found to affect their technology adoption preferences [33] and product attraction and avoidance [61]. Specifically, access to health care and health care utilization have been considered in prior studies as important predictors and controls of technology adoption in the consumer health context (eg, [41]). However, the direction of the influence of these situational factors on consumers' disposition to health care technologies and technology services channels, such as mHealth, is unclear. Frequent usage of

in-person health services may imply that strong relationships have been established with health providers and technology intermediation may only be considered when patients are dissatisfied with their providers (eg, [62]). However, a recent study suggested the opposite and those who had strong relationships with their providers were more likely to use personal health records [63]. Therefore, we consider variation in the following consumer-level situational characteristics specific to health care availability and utilization: (1) distance to health facilities (both primary care and specialty facilities) as a proxy for health care access and, (2) whether or not the last health checkup was recent as a proxy for recent routine health care utilization.

Independent Variables: Socioeconomic Status and Demographics

Although research results on the influence of socioeconomic status (SES) and demographic variables on innovation adoption are sometimes mixed (eg, [33]), in general, younger people [64], people with higher levels of education [65,66], and people with higher levels of income [65] are often found to be more innovative toward technology, including mobile services. However, mHealth also has the potential to be attractive to those who do not have a computer and Internet connection at home, but are still users of convenient mobile services (eg, [67,68]). Therefore, we account for the influence of SES and demographic characteristics on mHealth usage intentions, assimilation, and channel preferences by considering age, gender, income, and education in our research model.

Methods

Survey Design, Development, and Pretesting

Based on our research model, we designed a cross-sectional survey to measure consumers' mHealth usage intentions, mHealth assimilation, and preference to use mHealth as a complement or substitute to in-person doctor visits. In addition, data regarding PIMS, health care availability, health care utilization, health perceptions, SES, and demographics were collected. Existing instruments were applied whenever possible. All questions were adapted to the mHealth context. Extensive pretesting was conducted before final administration of the survey. We invited a total of 20 reviewers, including physicians, technologists, researchers, and managers working in or very familiar with the mHealth industry, to examine the survey instrument in detail before pilot-testing the survey. Although most of the expert feedback indicated that the questions were clear and easy to understand, necessary revisions were made according to their suggestions. After initial survey refinement, we conducted an online pilot study with 134 consumers in the United States to further assess the psychometric properties of the measures. Further refinements were made to the survey wording on the basis of pilot study results. A summary of final survey items and measures is available in [Multimedia Appendix 1](#).

Survey Sample, Recruitment, and Administration

To facilitate the data collection and administration process, we recruited a market research company. We worked closely with

this company to ensure that the sample was stratified to represent the US population in terms of age, gender, education, and income (following the US census) and that nonresponse bias was minimized. Using the online panel from the market research company, 8673 invitation emails were sent in 5 successive waves during a 2-week data collection period. In an effort to obtain a final sample that was nationally representative, we systematically monitored the demographics of incoming responses in each of the 5 waves and compared the means of the aggregate demographics to US census distributions. Oversampling was conducted in the subsequent waves for underrepresented strata (including those in younger and less-educated strata) to yield a final sample that was reasonably representative of the US census.

Each participant was provided with a unique passcode to access the online questionnaire. This design protected personal information from unauthorized access and also prevented duplicate responses from the same individual. Reminder emails were sent to participants to encourage them to complete the survey within the fieldwork period. The potential for nonresponse bias was mitigated by placing an emphasis on obtaining a nationally representative sample, following up with nonresponders and requesting participation, and including key demographic (age and gender) and socioeconomic (education and income) variables in the final models. Institutional Review Board (IRB) approval was obtained before survey administration. All participants acknowledged informed consent before taking the survey. Each participant took approximately 20 minutes to complete all 34 questions on the 17 screens of the online survey.

Analysis

Measurement quality of the multi-item measures was assessed through confirmatory factor analysis (CFA) and detailed assessment of construct reliability and validity. The details of the measurement quality analysis are available in the next section and further detail is available in [Multimedia Appendix 2](#). Final estimation for the primary models was completed using hierarchical ordinary least squares (OLS) regressions with robust standard error estimation. Several additional models and tests (eg, 2-stage least square analyses, mediated models with the demographic variables affecting the dependent variables through PIMS as well as directly, models with additional interactions)

were evaluated to examine the robustness of the OLS results. The results stood up to the robustness checks. The OLS estimations are reported for the primary results because of their straightforward interpretation. Secondly, we used ANOVA to assess differences in mHealth usage intentions and mHealth assimilation between respondents who preferred substitutive use of mHealth more strongly than complementary use versus those who preferred complementary use of mHealth more strongly than substitutive use.

Results

Descriptive Statistics

Complete data from 1132 respondents were collected. We carefully examined the distribution of respondents in our sample and found it to be nationally representative as compared with the distributions reported in the 2012 US Census [69] ([Table 1](#)). We obtained a final response rate of 13.05%, which is similar to response rates obtained in other online surveys conducted in comparable contexts (eg, [41,63]). Early- and late-stage respondents differed only by age and education. These differences were expected because of later-stage purposive oversampling of underrepresented strata. As such, the early-versus late-stage respondent analyses did not reveal any evidence of nonresponse bias. We conducted a marker variable analysis [70] and did not find evidence of common method bias.

The sample was relatively balanced in terms of gender (513 male and 619 female). The average age was 45 years (range 18-86, SD 16.20); 227 respondents (20.05%) were older than 60 years. Most respondents lived more than 6 miles from general and specialized health care facilities. Respondents had varying levels of education and individual income, representing reasonable variation in socioeconomic status. Additionally, 611 of 1132 respondents (53.98%) felt that they were healthy or very healthy, and there was substantial variance among respondents on the level of concern for vulnerability to chronic disease. Of the 1132 respondents, 430 (37.99%) reported that they had started using mHealth and 215 individuals (18.99%) reported use of mHealth on a regular basis. Further, 532 of 1132 respondents (47.00%) indicated that they would use mHealth as a substitute to in-person doctor visits, whereas 758 individuals (66.96%) indicated that they would use mHealth as a complement.

Table 1. Sample characteristics (N=1132).

Variables and categories	Sample, n (%)	US Census (%)
Demographics		
Age (years)		
18-29	218 (19.27)	22.1
30-39	269 (23.76)	17.1
40-49	169 (14.93)	18.6
50-59	249 (22.00)	17.9
60-69	155 (13.69)	11.8
≥70	72 (6.36)	12.5
Gender		
Male	513 (45.31)	49.2
Female	619 (54.68)	50.8
Socioeconomic status		
Education		
Not a high school graduate	18 (1.59)	12.9
High school graduate	211 (18.64)	31.2
Some college, but no degree	344 (30.39)	16.8
Associate's degree	154 (13.60)	9.1
Bachelor's degree	286 (25.27)	19.4
Advanced degree	119 (10.51)	1.5
Individual income (US \$)		
Less than 24,999	430 (37.99)	55.0
25,000-49,999	344 (30.39)	24.0
50,000-74,999	214 (18.90)	22.0
75,000-99,999	85 (7.50)	5.0
≥100,000	59 (5.21)	5.0
Health care availability		
Distance to primary health care facility		
<1 mile	34 (3.00)	—
1-5 miles	90 (7.95)	—
6-10 miles	472 (41.70)	—
≥11 miles	375 (33.13)	—
Do not know	161 (14.22)	—
Distance to specialized health care facility		
<1 mile	86 (7.60)	—
1-5 miles	57 (5.04)	—
6-10 miles	341 (30.12)	—
≥11 miles	381 (33.66)	—
Do not know	267 (23.59)	—
Health care utilization		
Recent health checkup		
No	125 (11.04)	—
Yes, with the past 5 years	37 (3.27)	—

Variables and categories	Sample, n (%)	US Census (%)
Yes, within the past 3 years	128 (11.31)	—
Yes, within the past 1 year	842 (74.38)	—

Measurement Quality

Before conducting hierarchical multivariate OLS regression analyses of the response data, we performed a series of checks to ensure the quality of the survey measures. [Multimedia Appendix 2](#) provides a summary of means, standard deviations, and correlations for all variables as well as reliability and validity measures for multi-item constructs (eg, composite reliabilities and average variances extracted). The CFA was performed using AMOS 7.0 to assess the measurement properties of the 4 multi-item constructs (behavioral usage intention, substitutive use, complementary use, PIMS) at the model and item levels [71]. The 4-factor model yielded an adequate model fit (comparative fit index=0.98, goodness-of-fit index=0.96, and standardized root mean square residual=0.03) [72]. The factor loadings for each indicator on its corresponding construct were greater than 0.70 and were significant at $P<.05$, thus supporting convergent validity. For each construct, the average variance extracted (AVE) was greater than 0.5, suggesting that the explained variance was more than the unexplained variance [73]. Additionally, the square root of the AVE for each construct was more than all its interconstruct correlations, thereby establishing discriminant validity [74]. In terms of reliability, Cronbach alphas and composite reliabilities were all greater than the recommended 0.70 level [75]. These results suggest that the measurement scales exhibit good psychometric properties.

Data Analysis

Multivariate OLS regressions were used to analyze the determinants of mHealth usage intentions, assimilation, and channel preferences. We evaluated 4 models per dependent variable in hierarchical fashion: (1) demographic and SES variables, (2) model 1 plus health variables (ie, distance to primary and specialized health care facilities, perceived healthiness, perceived vulnerability, and recency of health checkup), (3) model 2 plus PIMS, and (4) model 3 plus interaction effects. We also controlled for whether or not the respondent currently used mHealth (adopter or nonadopter dummy variable) to obtain generalizable results across the pooled sample of adopters and nonadopters.

mHealth Behavioral Usage Intention and Assimilation Results

[Table 2](#) (models A1-A4) reports behavioral usage intention regression results and [Table 3](#) (models B1-B4) reports assimilation regression results.

In models A1 and B1 (demographic and SES variables only), 40.31% and 23.04% of variation in behavioral usage intention and assimilation was explained, respectively. Older respondents were associated with a lower level of behavioral usage intention ($\beta=-0.02$, $P<.001$) and a lower level of assimilation ($\beta=-0.04$, $P<.001$). In addition, individual income had a significant positive association with both behavioral usage

intention ($\beta=0.12$, $P=.009$) and assimilation ($\beta=0.71$, $P<.001$). Moreover, level of education was negatively associated with assimilation ($\beta=-0.11$, $P=.02$), but not significantly associated with behavioral usage intention ($\beta=0.00$, $P=.97$). When controlling for differences between mHealth adopters and nonadopters with a dummy variable (adopter=1, nonadopter=0) in the behavioral usage intention models (A1-A4), we found the adopter group to have significantly increased intentions to continue using mHealth as compared to nonadopters' intentions to begin using mHealth.

In models A2 and B2, health care access, health care utilization, and perceived health condition variables were added to the models, resulting in 44.06% and 44.86% variance explained, respectively. Respondents who felt healthier were positively associated with behavioral usage intention ($\beta=0.30$, $P<.001$) and assimilation ($\beta=0.71$, $P<.001$). Respondents who felt more vulnerable to chronic disease were associated with stronger behavioral usage intention ($\beta=0.36$, $P<.001$) and stronger assimilation ($\beta=0.90$, $P<.001$). In addition, the recency of health checkup significantly increased the level of assimilation ($\beta=0.17$, $P=.001$), but was not significantly associated with behavioral usage intention ($\beta=0.04$, $P=.90$). Distance to primary and specialized facilities were not significant predictors of either behavioral usage intention (primary: $\beta=-0.05$, $P=.43$; specialized: $\beta=0.09$, $P=.08$) or assimilation (primary: $\beta=0.09$, $P=.14$; specialized: $\beta=0.01$, $P=.82$).

In models A3 and B3, PIMS was added, increasing the explained variance (ΔR^2) for behavioral usage intention by 15.55% and for assimilation by 8.47%. The significant positive coefficients indicate that PIMS was positively related to both behavioral usage intention ($\beta=1.11$, $P<.001$) and assimilation ($\beta=0.84$, $P<.001$).

In models A4 and B4, the interaction between PIMS and perceived healthiness (ie, PIMS*HLTH) and the interaction between PIMS and perceived vulnerability (ie, PIMS*VULN) were added, increasing the explained variance (ΔR^2) for behavioral usage intention by 0.09% and for assimilation by 7.08%. Although the main effects were significant (HLTH, VULN, and PIMS) predictors of both behavioral usage intention and assimilation, the interactions were not significant predictors of behavioral usage intention (PIMS*HLTH: $\beta=0.03$, $P=.41$; PIMS*VULN: $\beta=0.06$, $P=.10$), but were significant predictors of assimilation (PIMS*HLTH: $\beta=0.50$, $P<.001$; PIMS*VULN: $\beta=0.43$, $P<.001$). These results indicate that the main effects of PIMS, HLTH, and VULN are important factors that predict mHealth behavioral usage intention. Additionally, these factors not only independently, but also jointly, influence assimilation.

To develop a more nuanced understanding of the significant interaction effects in the mHealth assimilation model (B4), we plotted the interaction effects between PIMS and perceived

healthiness (PIMS*HLTH) as well as between PIMS and perceived chronic disease vulnerability (PIMS*VULN). We performed simple slope tests of the effects of HLTH and VULN on assimilation at different levels of the moderator (ie, PIMS) as recommended by Aiken and West [76]. We observed that (1) respondents with high PIMS reported higher levels of

assimilation when they reported higher health perceptions or higher perceived health vulnerability, whereas (2) respondents with low PIMS reported lower levels of assimilation than those with high PIMS, with the reported assimilation being even lower if respondents perceived themselves to be healthier or more vulnerable to chronic disease (Figures 2 and 3).

Table 2. Hierarchical ordinary least squares (OLS) regressions for consumer mHealth behavioral usage intention.

Variables	mHealth behavioral usage intention, OLS estimation (robust SE)			
	Model A1: Demographics and SES	Model A2: Health variables	Model A3: Personal innovativeness	Model A4: Interaction effects
Demographics				
Age (continuous in years)	-0.02 (0.00) ^c	-0.02 (0.00) ^c	0.00 (0.00)	0.00 (0.00)
Gender (female=1)	0.11 (0.10)	0.06 (0.09)	-0.00 (0.08)	-0.01 (0.08)
Socioeconomic status				
Education (5=Master's degree+)	0.00 (0.04)	0.05 (0.04)	0.03 (0.03)	0.03 (0.03)
Individual income (5≥US \$100K)	0.12 (0.04) ^b	0.03 (0.04)	-0.06 (0.04)	-0.06 (0.04)
Dummy				
Adopter (1)/nonadopter (0)	2.33 (0.10) ^c	1.97 (0.11) ^c	1.17 (0.11) ^c	1.14 (0.11) ^c
Health care availability				
Distance to primary facility	—	-0.05 (0.06)	-0.02 (0.05)	-0.03 (0.05)
Distance to specialized facility	—	0.09 (0.05)	0.08 (0.05)	0.08 (0.05)
Health care utilization				
Recent health checkup	—	0.04 (0.05)	-0.01 (0.04)	-0.01 (0.04)
Perceived health conditions				
Perceived healthiness (HLTH)	—	0.30 (0.05) ^c	0.12 (0.05) ^b	0.10 (0.05) ^a
Perceived vulnerability (VULN)	—	0.36 (0.05) ^c	0.18 (0.04) ^c	0.16 (0.05) ^b
Personal innovativeness toward mobile services (PIMS)				
PIMS	—	—	1.11 (0.06) ^c	1.11 (0.06) ^c
PIMS*HLTH	—	—	—	0.03 (0.04)
PIMS*VULN	—	—	—	0.06 (0.04)
Constant	3.95 (0.21) ^c	4.05 (0.24) ^c	3.53 (0.20) ^c	3.50 (0.21) ^c
R ²	0.4031	0.4406	0.5961	0.5970
ΔR ²	—	0.0375	0.1555	0.0009
F _{df} statistic	—	16.84 _{5,1121} ^c	358.13 _{1,1120} ^c	1.74 _{2,1118}

^aP<.05.

^bP<.01.

^cP<.001.

Table 3. Hierarchical ordinary least squares (OLS) regressions for consumer mHealth assimilation.

Variables	mHealth assimilation, OLS estimation (robust SE)			
	Model B1: Demographics and SES	Model B2: Health variables	Model B3: Personal innovativeness	Model B4: Interaction effects
Demographics				
Age (continuous in years)	-0.04 (0.00) ^c	-0.04 (0.00) ^c	-0.02 (0.00) ^c	-0.01 (0.00) ^c
Gender (female=1)	0.04 (0.12)	-0.10 (0.10)	-0.15 (0.09)	-0.16 (0.09)
Socioeconomic status				
Education (5=Master's degree+)	-0.11 (0.04) ^a	0.02 (0.04)	0.00 (0.04)	0.03 (0.03)
Individual Income (5≥US \$100K)	0.71 (0.06) ^c	0.38 (0.05) ^c	0.28 (0.05) ^c	0.24 (0.04) ^c
Health care availability				
Distance to primary facility	—	0.09 (0.06)	0.11 (0.06)	0.05 (0.05)
Distance to specialized facility	—	0.01 (0.05)	0.01 (0.05)	0.01 (0.04)
Health care utilization				
Recent health checkup	—	0.17 (0.05) ^b	0.10 (0.05) ^a	0.10 (0.04) [*]
Perceived health conditions				
Perceived healthiness (HLTH)	—	0.71 (0.06) ^c	0.51 (0.06) ^c	0.38 (0.06) ^c
Perceived vulnerability (VULN)	—	0.90 (0.05) ^c	0.69 (0.05) ^c	0.43 (0.05) ^c
Personal innovativeness toward mobile services (PIMS)				
PIMS	—	—	0.84 (0.57) ^c	0.85 (0.05) ^c
PIMS*HLTH	—	—	—	0.50 (0.05) ^c
PIMS*VULN	—	—	—	0.43 (0.04) ^c
Constant	4.09 (0.26) ^c	4.11 (0.25) ^c	3.27 (0.24) ^c	2.95 (0.23) ^c
R^2	0.2304	0.4486	0.5333	0.6041
ΔR^2	—	0.2182	0.0847	0.0708
F_{df} statistic	—	71.15 _{5,1122} ^c	219.25 _{1,1121} ^c	151.24 _{2,1119} ^c

^a $P < .05$.

^b $P < .01$.

^c $P < .001$.

Figure 2. Moderating effect of personal innovativeness toward mobile services (PIMS) on perceived healthiness for mHealth usage assimilation: Model B4 PIMS*HLTH.

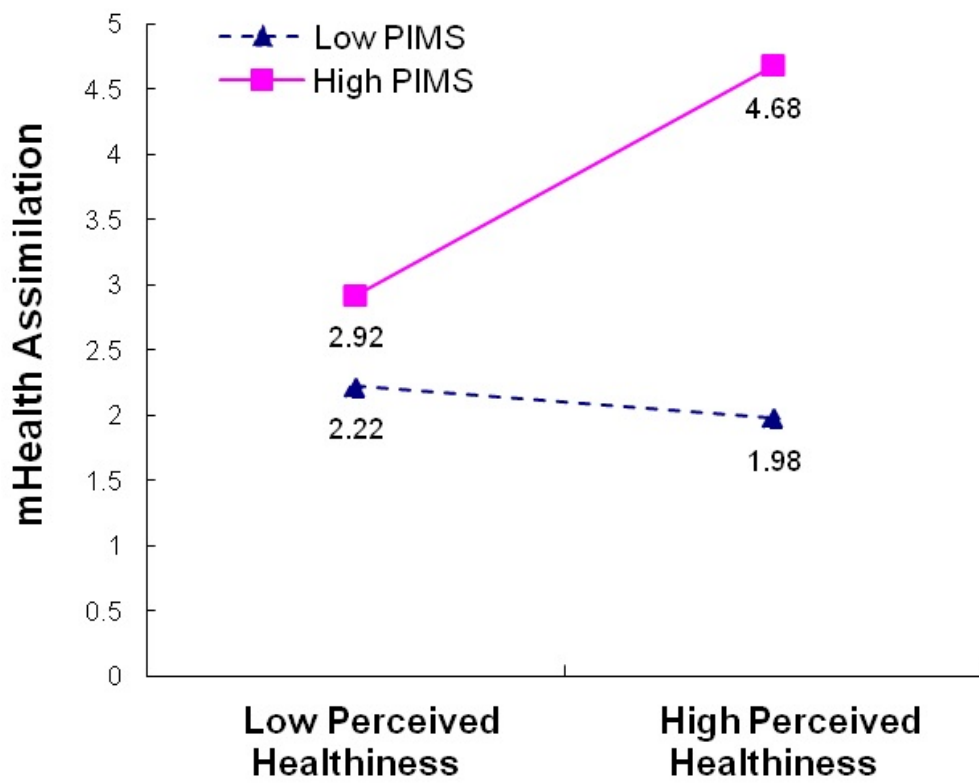
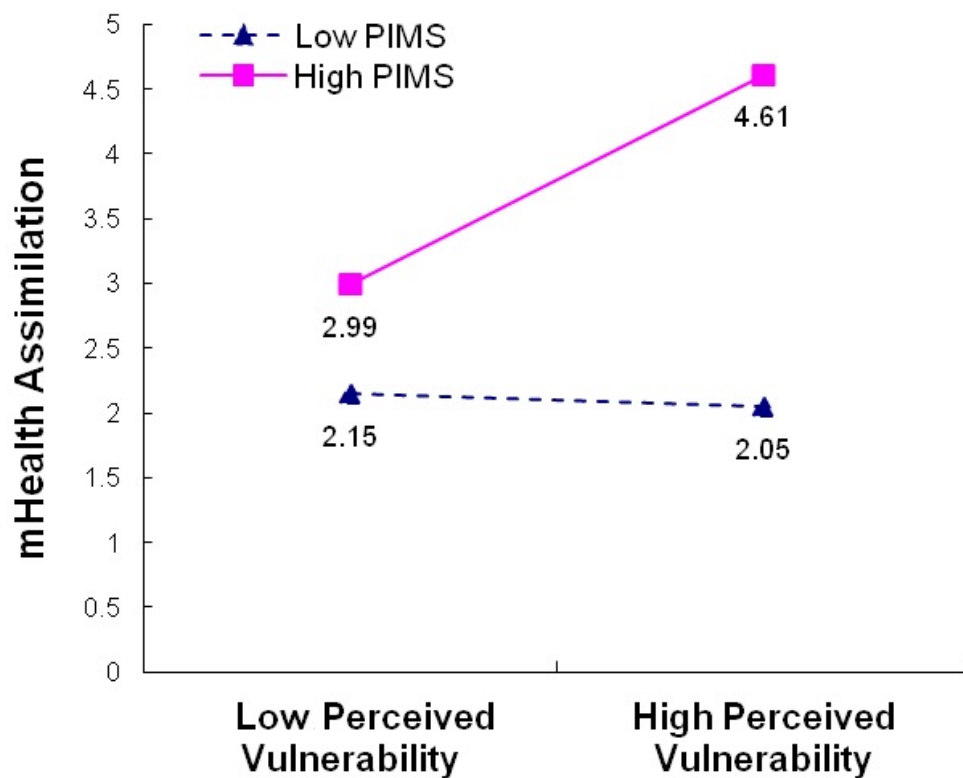


Figure 3. Moderating effect of personal innovativeness toward mobile services (PIMS) on perceived vulnerability for mHealth usage assimilation: Model B4 PIMS*VULN.



mHealth Substitutive and Complementary Use Preference Results

Table 4 (models C1-C4) reports substitutive use preference regression results, and **Table 5** (models D1-D4) reports complementary use preference regression results. In models C1 and D1, 27.48% and 18.43% of the variance was explained, respectively. Age was negatively associated with both substitutive use preference ($\beta=-0.03$, $P<.001$) and complementary use preference ($\beta=-0.02$, $P<.001$). In addition, individuals with higher levels of income were associated with higher substitutive use preference ($\beta=0.13$, $P=.003$), but not complementary use preference ($\beta=0.05$, $P=.23$). Moreover, females were more favorable to using mHealth as a complement to in-person doctor visits ($\beta=0.24$, $P=.009$) than males, yet there was not a significant gender difference for using mHealth as a substitute to in-person doctor visits ($\beta=0.15$, $P=.11$). The adopter group (adopter=1, nonadopter=0) was associated with higher substitutive use preference than nonadopters in all substitutive models (C1-C4) and in all complementary models (D1-D3), except for the interaction effects model (D4).

In models C2 and D2, health care access, health care utilization, and perceived health condition variables were added, increasing the explained variance (ΔR^2) by 6.91% for substitutive use preference and by 4.52% for complementary use preference. Perceived healthiness and perceived vulnerability were positively related to both substitutive use preference (perceived healthiness: $\beta=0.30$, $P<.001$; perceived vulnerability: $\beta=0.46$, $P<.001$) and complementary use preference (perceived healthiness: $\beta=0.20$, $P<.001$; perceived vulnerability: $\beta=0.34$, $P<.001$). Although a recent health checkup was not significantly associated with complementary use preference ($\beta=0.03$, $P=.56$), it was significantly negatively associated with substitutive use preference ($\beta=-0.16$, $P<.001$).

In models C3 and D3, PIMS was included, increasing explained variance (ΔR^2) by 9.56% for substitutive use preference and by 10.79% for complementary use preference. The positive and significant PIMS coefficients indicate that PIMS was a predictor of both substitutive use preference ($\beta=0.76$, $P<.001$) and complementary use preference ($\beta=0.75$, $P<.001$).

In models C4 and D4, the interaction between PIMS and perceived healthiness (ie, PIMS*HLTH) and the interaction between PIMS and perceived vulnerability (ie, PIMS*VULN) were added to the models, resulting in modest increases in explained variance (ΔR^2) by 1.35% for substitutive use preference and by 0.55% for complementary use preference. Both interaction terms were significant predictors for complementary use preference (PIMS*HLTH: $\beta=0.19$, $P<.001$; PIMS*VULN: $\beta=0.14$, $P<.001$), whereas only PIMS*HLTH was a significant predictor for substitutive use preference (PIMS*HLTH: $\beta=0.12$, $P=.02$; PIMS*VULN: $\beta=0.07$, $P=.11$). Overall, these results indicate that PIMS and

perceived health conditions jointly influence consumers' preferences between in-person doctor visits and mHealth.

We again plotted the interaction effects and conducted simple slope tests. For respondents with high PIMS, we observed a greater preference for mHealth as a substitute to in-person doctor visits when they felt healthier or more vulnerable to chronic disease (**Figures 4** and **5**). In contrast, for respondents with low PIMS, those who felt healthier were marginally less likely to use mHealth as a substitute for in-person doctor visits, but those who felt more vulnerable were more likely to prefer mHealth as a substitute.

As for the interaction effects associated with complementary use preference (**Figure 6**), respondents with high PIMS who felt healthier showed a greater preference for mHealth as a complement to in-person doctor visits relative to those who felt less healthy. For respondents with low PIMS, those who felt healthier indicated a marginally weaker preference for using mHealth as a complement relative to those who felt less healthy. The interaction between PIMS and VULN (PIMS*VULN) was not significant for model D4; thus, it is not presented here as a graph.

Given that the correlation of complementary use and substitutive use is 0.73, the preferences for complementary and substitutive use of mHealth can be interpreted to be mutually reinforcing. To further explore differences in those with stronger preferences for complementary use of mHealth than substitutive use of mHealth, we conducted an ANOVA analysis of differences in behavioral usage intention and assimilation between the following 2 groups: (1) stronger preference for complementary use than substitutive use (complementary > substitutive), and (2) stronger preference for substitutive use than complementary use (substitutive > complementary). We found that behavioral usage intention was significantly higher for the substitutive > complementary group (behavioral usage intention mean 4.69, SD 1.68) than for the complementary > substitutive group (behavioral usage intention mean 3.92, SD 1.79; $F_{1,702}=20.14$, $P<.001$). Similarly, assimilation was significantly higher for the substitutive > complementary group (assimilation mean 3.78, SD 2.21) than for the complementary > substitutive group (assimilation mean 2.66, SD 1.66; $F_{1,702}=41.866$, $P<.001$). Perceived health and perceived vulnerability to chronic disease were not significantly different between the 2 groups. These results reveal that although more respondents in our sample were willing to use mHealth as a complement than as a substitute for in-person doctor visits, the substitutive > complementary group indicated stronger behavioral usage intention and assimilation than the complementary > substitutive group. Additionally, respondents in the substitutive > complementary group, relative to the complementary > substitutive group, were younger (mean age 39.00 years, SD 14.20 vs 46.35 years, SD 16.81; $F_{1,702}=21.284$, $P<.001$), had higher PIMS (PIMS mean 4.72, SD 1.68 vs 3.96, SD 1.76; $F_{1,702}=19.699$, $P<.001$), and were less likely to be female (46%, SD 0.50 vs 55%, SD 0.50; $F_{1,702}=4.033$, $P=.045$).

Table 4. Hierarchical ordinary least squares (OLS) regressions for mHealth substitutive use preference.

Variables	mHealth substitutive use preference, OLS estimation (robust SE)			
	Model C1: Demographics and SES	Model C2: Health variables	Model C3: Personal innovativeness	Model C4: Interaction effects
Demographics				
Age (continuous in years)	-0.03 (0.00) ^c	-0.03 (0.00) ^c	-0.01 (0.00) ^b	-0.01 (0.00) ^b
Gender (female=1)	0.15 (0.09)	0.14 (0.09)	0.10 (0.08)	0.09 (0.08)
Socioeconomic status				
Education (5=Master's degree+)	-0.05 (0.04)	0.01 (0.03)	-0.00 (0.03)	0.00 (0.03)
Individual Income (5≥US \$100K)	0.13 (0.04) ^b	0.06 (0.04)	-0.00 (0.04)	-0.01 (0.04)
Dummy				
Adopter (1)/nonadopter(0)	1.42 (0.11) ^c	1.06 (0.11) ^c	0.50 (0.11) ^c	0.40 (0.11) ^c
Health care availability				
Distance to primary facility	—	0.01 (0.06)	0.03 (0.05)	0.01 (0.05)
Distance to specialized facility	—	0.03 (0.05)	0.02 (0.04)	0.02 (0.04)
Health care utilization				
Recent health checkup	—	-0.16 (0.04) ^c	-0.20 (0.04) ^c	-0.19 (0.04) ^c
Perceived health conditions				
Perceived healthiness (HLTH)	—	0.30 (0.05) ^c	0.18 (0.05) ^c	0.15 (0.05) ^b
Perceived vulnerability (VULN)	—	0.46 (0.05) ^c	0.34 (0.05) ^c	0.26 (0.05) ^c
Personal innovativeness toward mobile services (PIMS)				
PIMS	—	—	0.76 (0.06) ^c	0.78 (0.06) ^c
PIMS*HLTH	—	—	—	0.19 (0.05) ^c
PIMS*VULN	—	—	—	0.14 (0.04) ^c
Constant	4.72 (0.21) ^c	5.15 (0.22) ^c	4.79 (0.21) ^c	4.71 (0.21) ^c
R^2	0.2748	0.3439	0.4395	0.4530
ΔR^2	—	0.0691	0.0956	0.0135
F_{df} statistic	—	24.91 _{5,1121} ^c	171.53 _{1,1120} ^c	14.12 _{2,1118} ^b

^a $P < .05$.^b $P < .01$.^c $P < .001$.

Table 5. Hierarchical ordinary least squares (OLS) regressions for mHealth complementary use preference.

Variables	mHealth complementary use preference, OLS estimation (robust SE)			
	Model D1: Demographics and SES	Model D2: Health variables	Model D3: Personal innovativeness	Model D4: Interaction effects
Demographics				
Age (continuous in years)	-0.02 (0.00) ^c	-0.02 (0.00) ^c	-0.01 (0.00) ^a	-0.01 (0.00) ^a
Gender (female=1)	0.24 (0.09) ^b	0.20 (0.09) ^a	0.16 (0.08)	0.15 (0.08)
Socioeconomic status				
Education (5=Master's degree+)	0.03 (0.03)	0.08 (0.03) ^a	0.07 (0.03) ^a	0.07 (0.03) ^a
Individual income (5≥US \$100K)	0.05 (0.04)	-0.02 (0.04)	-0.07 (0.04)	-0.08 (0.04) ^a
Dummy				
Adopter (1)/nonadopter (0)	1.07 (0.09) ^c	0.76 (0.10) ^c	0.21 (0.10) ^a	0.15 (0.10)
Health care availability				
Distance to primary facility	—	0.03 (0.06)	0.05 (0.05)	0.03 (0.05)
Distance to specialized facility	—	0.05 (0.05)	0.05 (0.04)	0.05 (0.04) ^a
Health care utilization				
Recent health checkup	—	0.03 (0.04)	-0.01 (0.04)	-0.01 (0.04)
Perceived health conditions				
Perceived healthiness (HLTH)	—	0.20 (0.05) ^c	0.09 (0.05)	0.07 (0.05)
Perceived vulnerability (VULN)	—	0.34 (0.05) ^c	0.22 (0.04) ^c	0.17 (0.05) ^b
Personal innovativeness toward mobile services (PIMS)				
PIMS	—	—	0.75 (0.06) ^c	0.76 (0.06) ^c
PIMS*HLTH	—	—	—	0.12 (0.05) ^a
PIMS*VULN	—	—	—	0.07 (0.05)
Constant	5.09 (0.19) ^c	5.09 (0.22) ^c	4.73 (0.21) ^c	4.69 (0.21) ^c
R^2	0.1843	0.2295	0.3374	0.3429
ΔR^2	—	0.0452	0.1079	0.0055
F_{df} statistic	—	15.43 _{5,1121} ^c	140.90 _{1,1120} ^c	4.00 _{2,1118} ^a

^a $P < .05$.^b $P < .01$.^c $P < .001$.

Figure 4. Moderating effect of personal innovativeness toward mobile services (PIMS) on perceived healthiness for preferring mHealth as a substitute to doctor visits: Model C4 PIMS*HLTH.

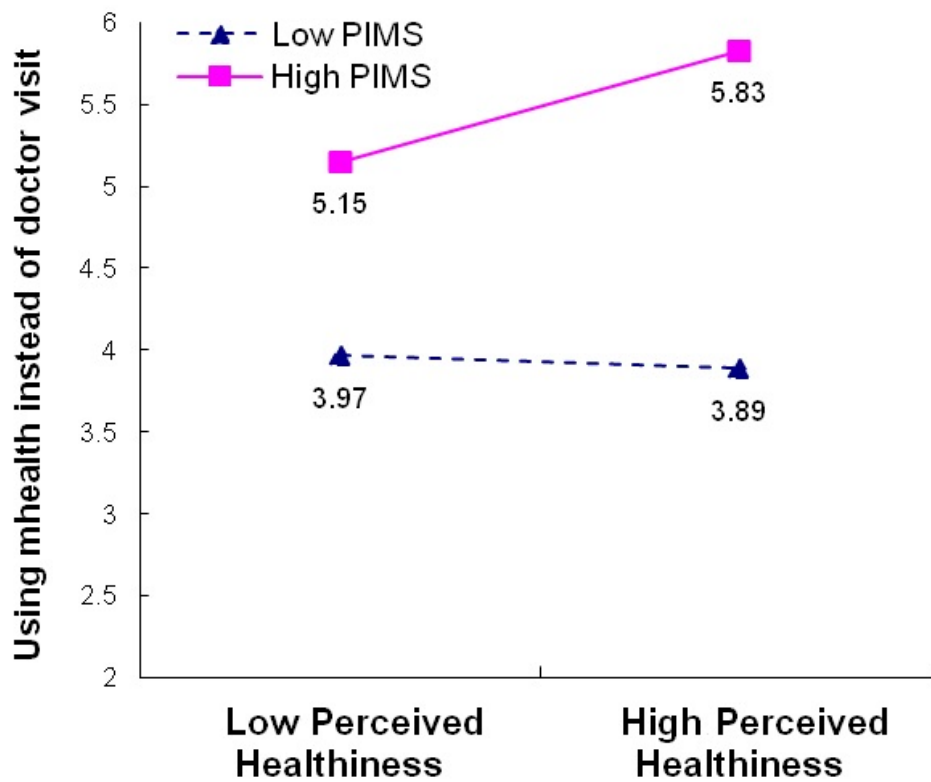


Figure 5. Moderating effect of personal innovativeness toward mobile services (PIMS) on perceived vulnerability for preferring mHealth as a substitute to doctor visits: Model C4 PIMS*VULN.

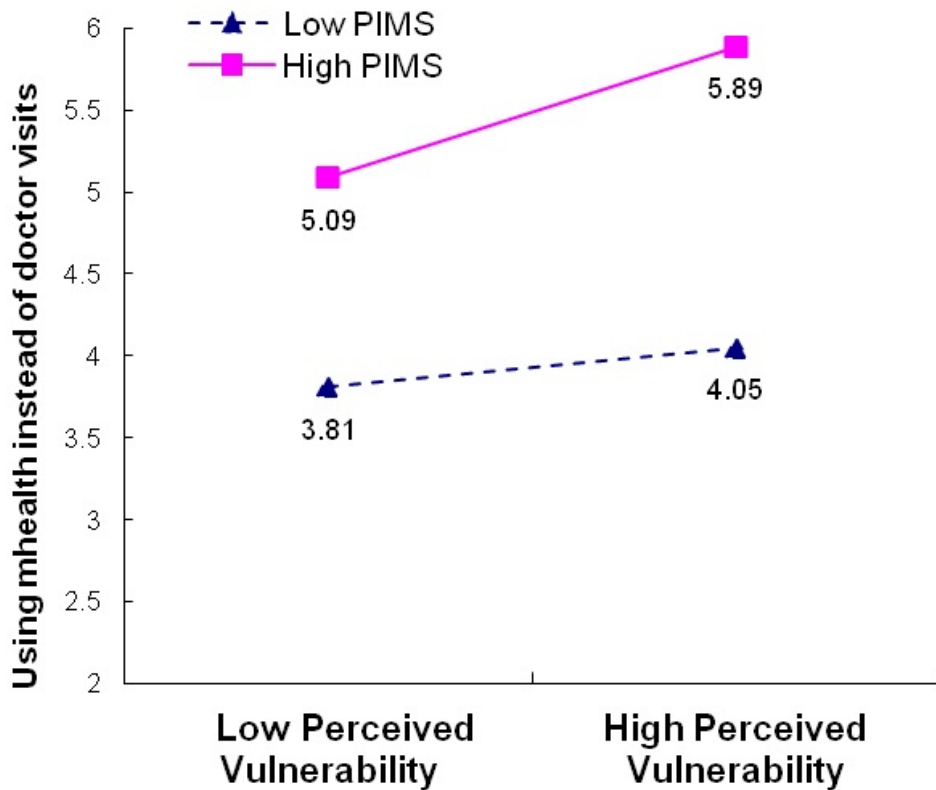
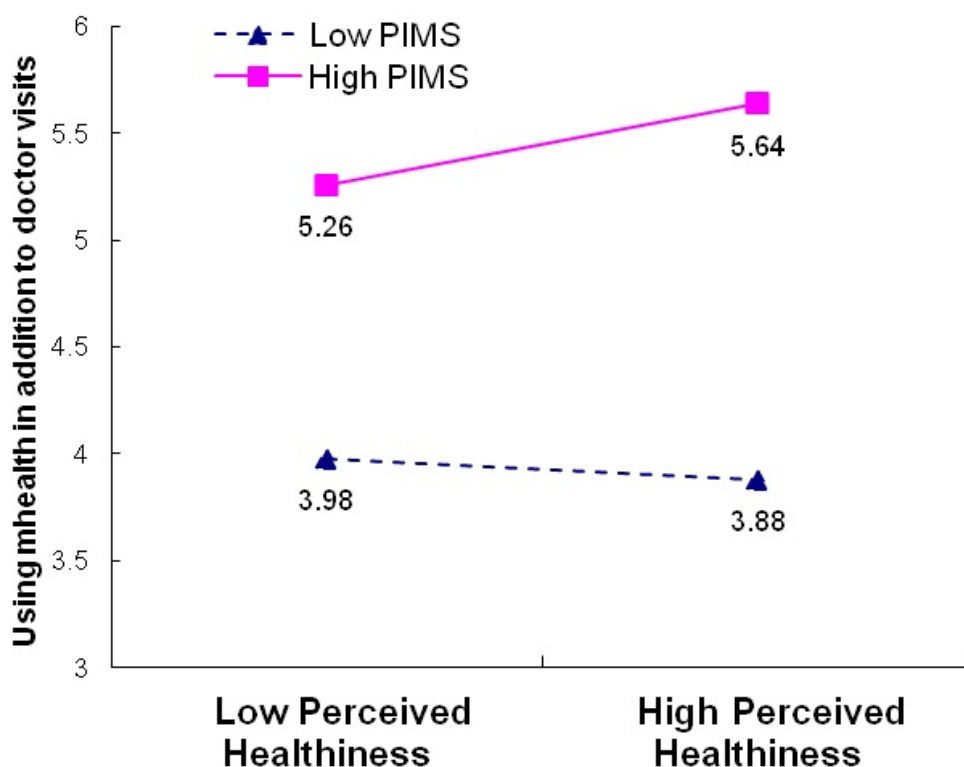


Figure 6. Moderating effect of personal innovativeness toward mobile services (PIMS) on perceived healthiness for preferring mHealth as a complement to doctor visits: Model D4 PIMS*HLTH.



Discussion

Principal Findings and Implications

The market for mHealth is growing rapidly, but research in this emerging area has been limited. Before this study, there was limited understanding as to what determinants were associated with mHealth usage intentions, extent of mHealth assimilation, and why mHealth might be preferred as a complementary or substitutive service consumption channel in a context traditionally associated with hands-on, in-person interactions. This study has provided novel insights by examining how consumer usage intentions and assimilation of mHealth, as well as consumer channel preferences for health services are affected by PIMS, perceived health conditions, health care availability, health care utilization, demographics, and socioeconomic status.

Our primary findings are as follows: (1) more consumers are favorable to using mHealth as a complement to in-person doctor visits than as a substitute, but those who prefer mHealth as a substitute report stronger usage intentions and higher assimilation than those who prefer mHealth as a complement although not being significantly different in health perceptions, and (2) PIMS and perceived health conditions have direct effects on usage intentions, assimilation, and channel preferences and mutually reinforcing effects on assimilation and channel preferences. Of particular interest is the finding that the combination of higher PIMS and increased perceptions of healthiness as well as the combination of higher PIMS and increased perceptions of chronic disease vulnerability are significantly associated with higher mHealth assimilation and substitutive use of mHealth. The combination of higher PIMS

and increased perceptions of healthiness is also significantly positively associated with the complementary use of mHealth. These interesting findings suggest that current health status is not the only predictor of mHealth usage and, perhaps counterintuitively, it is not necessarily those who perceive themselves as the least healthy who are the most likely to adopt and use mHealth.

We also demonstrate that PIMS and perceived chronic disease vulnerability are important positive predictors. Given these results, individuals worried about diet, weight, blood pressure, exercise, and other health issues might consider proactively using apps such as MyFitnessPal [77], InstantHeartRate [78], Macaw [79], and Livestrong [14]. Such proactive management of one's health could significantly reduce the incidence of chronic disease and reduce the burden of such conditions on our health system [80].

Additionally, our results demonstrate that more than one-third of respondents specified current use of mHealth, whereas almost one-fifth of respondents report currently using mHealth on a regular basis. In addition, approximately two-thirds of respondents said they would use mHealth as a complement to in-person doctor visits. These findings elaborate prior research suggesting that individual innovativeness [50], individual traits [32], and health self-perceptions [81] are associated with usage intentions. Our findings extend prior research by considering the influence of these constructs, including the interaction effects of perceived health conditions, on consumers' intentions, assimilation, and channel preferences for mHealth use. However, as suggested by the IT-enabled self-service literature, the infusion of technology into a service encounter may be met with resistance by those who prefer a hands-on relationship (vs a

high-tech relationship) [46]. Our results support such findings in the mHealth context, in that health care utilization (ie, recent health checkup) had a positive effect on mHealth assimilation, but a negative effect on using mHealth as a substitute to in-person doctor visits. These results may indicate that personal relationships in health care settings will be difficult to augment (or replace) with technology for certain consumer segments. Although other technologies, such as telemedicine, have provided mechanisms to extend health services to those with limited access (eg, [82]), the issue of how mHealth can improve health care access while not adversely affecting patient-provider relationships will be an essential consideration.

Finally, we find that increased age is associated with decreased usage intentions and assimilation of mHealth in many of our models, whereas increased income is associated with increased usage intentions and usage of mHealth in some models. Similar findings have been reported in other technology acceptance studies [33,64,65]. Additionally, somewhat contrary to prior research suggesting that higher levels of education are often positively associated with technology adoption [65,66], we find a mix of significant and nonsignificant effects of education in our models. We did find that education was a positive and significant predictor within many of our models associated with using mHealth as a complement to in-person doctor's office visits. This could be an area for further research.

Strengths and Limitations

The strengths of our study include: (1) a theoretically driven model, based on technology acceptance, technology assimilation, consumer behavior, and health informatics literature, on the determinants of consumer mHealth usage intentions, assimilation, and channel preferences, (2) the inclusion of direct and interactive (moderating) effects of PIMS and perceived health conditions (vulnerability and healthiness) as determinants, and (3) robust survey, sampling, and analysis methods. Our study is limited by the cross-sectional nature of our survey. We

note that our robustness checks included 2-stage estimation models and mediated models. All our findings held up to these checks, but future research could consider longitudinal research designs to elaborate our understanding of the mechanisms through which usage intentions and assimilation of mHealth develop. We are also limited by the use of an online survey, which may be biased toward those who complete surveys online or respondents who are more technologically sophisticated. Future research could consider other surveying and sampling strategies. Although our models have the feature of parsimony, they may exclude other situational, demographic, or individual characteristics. Future research could expand upon our findings by including such additional characteristics. Finally, our results are generalizable to the general population because the chosen sampling strategy and the use of statistical controls. However, future research could delve deeper into subgroup differences (adopters vs nonadopters, health respondents vs unhealthy respondents, resource-rich respondents vs resource-poor respondents, etc) and provide more nuanced findings regarding between and within group heterogeneity.

Conclusions

This study has provided insights into the usage intentions, assimilation, and channel preferences associated with mHealth. These findings contribute to the health informatics literature and to health policy initiatives by demonstrating that mHealth will face both acceptance and resistance. Targeting the most receptive consumer segments may be the best strategy to encourage widespread diffusion. Multiple predictors have been shown to have significant impacts on mHealth preferences and both direct and interactive effects were observed. We suggest that future initiatives to promote mHealth should shift targeting of consumers from coarse demographics to individual dispositions toward mobile service innovations, complementary or substitutive channel use preferences, and perceived health conditions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Constructs and measurement sources.

[[PDF File \(Adobe PDF File\), 290KB - jmir_v15i8e149_app1.pdf](#)]

Multimedia Appendix 2

Reliability, average variance extracted, and correlations.

[[PDF File \(Adobe PDF File\), 389KB - jmir_v15i8e149_app2.pdf](#)]

References

1. Lester RT, van der Kop M, Taylor D, Alasaly K, Coleman J, Marra F. m-Health: Connecting patients to improve population and public health. *BCMJ* 2011;53(5):218-219 [[FREE Full text](#)]
2. Fox S, Duggan M. Mobile health. Washington, DC: Pew Internet & American Life Project; 2012 Nov 08. URL: http://www.pewinternet.org/~media/Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2013-03-12] [[WebCite Cache ID 6F4tz8FH7](#)]

3. PriceWaterhouseCoopers. Emerging mHealth: Paths for Growth. 2012. URL: http://www.pwc.com/en_GX/gx/healthcare/mhealth/assets/pwc-emerging-mhealth-exec-summary.pdf [accessed 2013-03-12] [WebCite Cache ID 6F4uIEZu2]
4. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *Lancet* 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: [10.1016/S0140-6736\(11\)60701-0](https://doi.org/10.1016/S0140-6736(11)60701-0)] [Medline: [21722952](https://pubmed.ncbi.nlm.nih.gov/21722952/)]
5. Eysenbach G. Consumer health informatics. *BMJ* 2000 Jun 24;320(7251):1713-1716 [FREE Full text] [Medline: [10864552](https://pubmed.ncbi.nlm.nih.gov/10864552/)]
6. Center for Health Market Innovations. 2004. Community Health Information Tracking System (CHITS) URL: <http://healthmarketinnovations.org/program/community-health-information-tracking-system-chits> [accessed 2013-03-12] [WebCite Cache ID 6F4uRVf9a]
7. Center for Health Market Innovations. 2005. Treatment and Research AIDS Center (TRAC) URL: <http://healthmarketinnovations.org/program/tracnet> [accessed 2013-03-12] [WebCite Cache ID 6F4uWliVs]
8. Center for Health Market Innovations. 2003. Alerta DISAMAR URL: <http://healthmarketinnovations.org/program/alerta-disamar> [accessed 2013-03-12] [WebCite Cache ID 6F4uMbr2G]
9. WebMD. iTunes. 2013. WebMD the Magazine URL: <https://itunes.apple.com/us/app/webmd-the-magazine/id503356717?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2ETbxHT]
10. SigmaPhone LLC. iTunes. 2012. iPharmacy - Pill ID & Rx Reminder URL: <https://itunes.apple.com/us/app/i-pharmacy-pill-id-rx-reminder/id348702163?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2Eflb9e]
11. Mayo Clinic. iTunes. 2013. Patient URL: <https://itunes.apple.com/us/app/patient/id523220194?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2EjRikW]
12. eClinicalWorks. 2013. eClinicalMobile URL: <http://www.eclinicalworks.com/products-eclinical-mobile.htm> [accessed 2013-07-11] [WebCite Cache ID 6I2EmLHBK]
13. sanofi-aventis US LLC. iTunes. 2013. GoMeals URL: <https://itunes.apple.com/us/app/gomeals/id336651139?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2Erok0r]
14. Livestrong.com. iTunes. 2013. Livestrong.com - Calorie Tracker - Your Diet and Fitness Calorie Counter for Better Health URL: <https://itunes.apple.com/us/app/livestrong.com-calorie-tracker/id295305241?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2ExoUzn]
15. WellDoc, Inc. iTunes. 2013. AT&T mHealth URL: <https://itunes.apple.com/us/app/at-t-mhealth/id480737917?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2F2pjG]
16. Azumio Inc. iTunes. 2013. Glucose Buddy - Diabetes Logbook Manager w/syncing, Blood Pressure, Weight Tracking URL: <https://itunes.apple.com/us/app/glucose-buddy-diabetes-logbook/id294754639?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2F6a8dj]
17. Cocosila M, Archer N. An empirical investigation of mobile health adoption in preventive interventions. In: BLED 2009 PROCEEDINGS. 2009 Presented at: 22nd Bled eConference; June 14-17, 2009; Bled, Slovenia.
18. Hung MC, Jen WY. The adoption of mobile health management services: an empirical study. *J Med Syst* 2012 Jun;36(3):1381-1388. [doi: [10.1007/s10916-010-9600-2](https://doi.org/10.1007/s10916-010-9600-2)] [Medline: [20878452](https://pubmed.ncbi.nlm.nih.gov/20878452/)]
19. Weiner JP. Doctor-patient communication in the e-health era. *Isr J Health Policy Res* 2012;1(1):33 [FREE Full text] [doi: [10.1186/2045-4015-1-33](https://doi.org/10.1186/2045-4015-1-33)] [Medline: [22929000](https://pubmed.ncbi.nlm.nih.gov/22929000/)]
20. 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/)]
21. Ohno-Machado L. Informatics 2.0: implications of social media, mobile health, and patient-reported outcomes for healthcare and individual privacy. *J Am Med Inform Assoc* 2012;19(5):683. [doi: [10.1136/amiajn-2012-001224](https://doi.org/10.1136/amiajn-2012-001224)] [Medline: [22879365](https://pubmed.ncbi.nlm.nih.gov/22879365/)]
22. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. *Am J Prev Med* 2008 Aug;35(2):177-181 [FREE Full text] [doi: [10.1016/j.amepre.2008.05.001](https://doi.org/10.1016/j.amepre.2008.05.001)] [Medline: [18550322](https://pubmed.ncbi.nlm.nih.gov/18550322/)]
23. Winstead-Derlega C, Rafaly M, Delgado S, Freeman J, Cutitta K, Miles T, et al. A pilot study of delivering peer health messages in an HIV clinic via mobile media. *Telemed J E Health* 2012;18(6):464-469. [doi: [10.1089/tmj.2011.0236](https://doi.org/10.1089/tmj.2011.0236)] [Medline: [22732025](https://pubmed.ncbi.nlm.nih.gov/22732025/)]
24. Baron J, Newman S. A systematic review of the effectiveness of mobile health interventions for the management of diabetes. London, UK: Igitur Publishing; 2011 Presented at: International Congress on Telehealth and Telecare; March 1-3, 2011; London, UK.
25. Schweitzer J, Synowiec C. The economics of eHealth and mHealth. *J Health Commun* 2012;17 Suppl 1:73-81. [doi: [10.1080/10810730.2011.649158](https://doi.org/10.1080/10810730.2011.649158)] [Medline: [22548602](https://pubmed.ncbi.nlm.nih.gov/22548602/)]
26. Curioso WH, Mechael PN. Enhancing 'M-health' with south-to-south collaborations. *Health Aff (Millwood)* 2010 Feb;29(2):264-267 [FREE Full text] [doi: [10.1377/hlthaff.2009.1057](https://doi.org/10.1377/hlthaff.2009.1057)] [Medline: [20348071](https://pubmed.ncbi.nlm.nih.gov/20348071/)]
27. Kahn JG, Yang JS, Kahn JS. 'Mobile' health needs and opportunities in developing countries. *Health Aff (Millwood)* 2010 Feb;29(2):252-258 [FREE Full text] [doi: [10.1377/hlthaff.2009.0965](https://doi.org/10.1377/hlthaff.2009.0965)] [Medline: [20348069](https://pubmed.ncbi.nlm.nih.gov/20348069/)]
28. Demiris G. New era for the consumer health informatics research agenda. *Health Syst* 2012 May 04;1(1):13-16. [doi: [10.1057/hs.2012.7](https://doi.org/10.1057/hs.2012.7)]

29. 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/)]
30. Hsieh J, Rai A, Keil M. Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. *MIS quarterly* 2008;32(1):97-126.
31. Agarwal R, Prasad J. A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research* 1998 Jun 01;9(2):204-215. [doi: [10.1287/isre.9.2.204](https://doi.org/10.1287/isre.9.2.204)]
32. Agarwal R, Prasad J. Are individual differences germane to the acceptance of new information technologies? *Decision Sciences* 1999 Mar;30(2):361-391. [doi: [10.1111/j.1540-5915.1999.tb01614.x](https://doi.org/10.1111/j.1540-5915.1999.tb01614.x)]
33. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009;16(4):550-560 [FREE Full text] [doi: [10.1197/jamia.M2888](https://doi.org/10.1197/jamia.M2888)] [Medline: [19390112](https://pubmed.ncbi.nlm.nih.gov/19390112/)]
34. Thompson TG, Brailer DJ. The Decade of Health Information Technology: Delivering Consumer-Centric and Information-Rich Health Care. 2004 Jul 21. URL: http://www.providersedge.com/ehdocs/ehr_articles/the_decade_of_hit-delivering_customer-centric_and_info-rich_hc.pdf [accessed 2013-03-26] [WebCite Cache ID [6FPQnkVPE](https://www.webcitation.org/6FPQnkVPE)]
35. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Management Science* 1989 Aug 01;35(8):982-1003. [doi: [10.1287/mnsc.35.8.982](https://doi.org/10.1287/mnsc.35.8.982)]
36. Fichman RG, Kemerer CF. The assimilation of software process innovations: an organizational learning perspective. *Management Science* 1997 Oct 01;43(10):1345-1363. [doi: [10.1287/mnsc.43.10.1345](https://doi.org/10.1287/mnsc.43.10.1345)]
37. Rai A, Brown P, Tang X. Organizational assimilation of electronic procurement innovations. *Journal of Management Information Systems* 2009 Aug 13;26(1):257-296. [doi: [10.2753/MIS0742-1222260110](https://doi.org/10.2753/MIS0742-1222260110)]
38. Dholakia UM, Kahn BE, Reeves R, Rindfleisch A, Stewart D, Taylor E. Consumer behavior in a multichannel, multimedia retailing environment. *Journal of Interactive Marketing* 2010 May;24(2):86-95. [doi: [10.1016/j.intmar.2010.02.005](https://doi.org/10.1016/j.intmar.2010.02.005)]
39. Nysveen H. Mobilizing the brand: the effects of mobile services on brand relationships and main channel use. *Journal of Service Research* 2005 Feb 01;7(3):257-276. [doi: [10.1177/1094670504271151](https://doi.org/10.1177/1094670504271151)]
40. Becker M. The health belief model and sick role behavior. *Health Educ Monogr* 1974;2:409-419.
41. Bethell C, Fiorillo J, Lansky D, Hendryx M, Knickman J. Online consumer surveys as a methodology for assessing the quality of the United States health care system. *J Med Internet Res* 2004 Jan 20;6(1):e2 [FREE Full text] [doi: [10.2196/jmir.6.1.e2](https://doi.org/10.2196/jmir.6.1.e2)] [Medline: [15111268](https://pubmed.ncbi.nlm.nih.gov/15111268/)]
42. Janz NK, Becker MH. The Health Belief Model: a decade later. *Health Educ Q* 1984;11(1):1-47. [Medline: [6392204](https://pubmed.ncbi.nlm.nih.gov/6392204/)]
43. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly* 2012;36(1):157-178.
44. Fichman RG. The diffusion assimilation of information technology innovations. In: Zmud RW, editor. *Framing the Domains of IT Management: Projecting the Future...Through the Past*. Cincinnati, OH: Pinnaflex Educational Resources Inc; 2000.
45. Berry LL, Seiders K, Grewal D. Understanding service convenience. *Journal of Marketing* 2002 Jul;66(3):1-17. [doi: [10.1509/jmkg.66.3.1.18505](https://doi.org/10.1509/jmkg.66.3.1.18505)]
46. Bitner MJ, Brown SW, Meuter ML. Technology infusion in service encounters. *Journal of the Academy of Marketing Science* 2000 Jan 01;28(1):138-149. [doi: [10.1177/0092070300281013](https://doi.org/10.1177/0092070300281013)]
47. Keh HT, Pang J. Customer reactions to service separation. *Journal of Marketing* 2010 Mar;74(2):55-70. [doi: [10.1509/jmkg.74.2.55](https://doi.org/10.1509/jmkg.74.2.55)]
48. Agarwal R, Prasad J. The antecedents and consequents of user perceptions in information technology adoption. *Decision Support Systems* 1998 Jan;22(1):15-29. [doi: [10.1016/S0167-9236\(97\)00006-7](https://doi.org/10.1016/S0167-9236(97)00006-7)]
49. Klein R. An empirical examination of patient-physician portal acceptance. *Eur J Inf Syst* 2007 Dec;16(6):751-760. [doi: [10.1057/palgrave.ejis.3000719](https://doi.org/10.1057/palgrave.ejis.3000719)]
50. Lu J, Yao JE, Yu C. Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. *The Journal of Strategic Information Systems* 2005 Sep;14(3):245-268. [doi: [10.1016/j.jsis.2005.07.003](https://doi.org/10.1016/j.jsis.2005.07.003)]
51. Thatcher JB, Perrewe PL. An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *MIS Quarterly* 2002;26(4):381-396.
52. Agarwal R, Sambamurthy V, Stair RM. Research report: the evolving relationship between general and specific computer self-efficacy--an empirical assessment. *Information Systems Research* 2000 Dec;11(4):418-430. [doi: [10.1287/isre.11.4.418.11876](https://doi.org/10.1287/isre.11.4.418.11876)]
53. Goldsmith RE. Explaining and predicting consumer intention to purchase over the internet: an exploratory study. *Journal of Marketing Theory and Practice* 2002;10(2):22-28.
54. Goldsmith RE. Using the Domain Specific Innovativeness Scale to identify innovative Internet consumers. *Internet Research* 2001;11(2):149-158. [doi: [10.1108/10662240110695098](https://doi.org/10.1108/10662240110695098)]
55. Citrin AV, Sprott DE, Silverman SN, Stem DE, Jr. Adoption of Internet shopping: the role of consumer innovativeness. *Industrial Management & Data Systems* 2000;100(7):294-300. [doi: [10.1108/02635570010304806](https://doi.org/10.1108/02635570010304806)]
56. Kuo Y, Yen S. Towards an understanding of the behavioral intention to use 3G mobile value-added services. *Computers in Human Behavior* 2009 Jan;25(1):103-110. [doi: [10.1016/j.chb.2008.07.007](https://doi.org/10.1016/j.chb.2008.07.007)]

57. Lu J, Yao JE, Yu C. Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. *The Journal of Strategic Information Systems* 2005 Sep;14(3):245-268. [doi: [10.1016/j.jsis.2005.07.003](https://doi.org/10.1016/j.jsis.2005.07.003)]
58. Moore GC, Benbasat I. Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research* 1991 Sep 01;2(3):192-222. [doi: [10.1287/isre.2.3.192](https://doi.org/10.1287/isre.2.3.192)]
59. Plouffe CR, Vandenbosch M, Hulland J. Intermediating technologies and multi-group adoption: A comparison of consumer and merchant adoption intentions toward a new electronic payment system. *Journal of Product Innovation Management* 2001 Mar;18(2):65-81. [doi: [10.1111/1540-5885.1820065](https://doi.org/10.1111/1540-5885.1820065)]
60. Rosenstock IM. Why people use health services. *Milbank Quarterly* 2005 Dec;83(4):Online-only. [doi: [10.1111/j.1468-0009.2005.00425.x](https://doi.org/10.1111/j.1468-0009.2005.00425.x)]
61. Bloch PH. Seeking the ideal form: product design and consumer response. *Journal of Marketing* 1995 Jul;59(3):16-29. [doi: [10.2307/1252116](https://doi.org/10.2307/1252116)]
62. Emont S. Measuring the impact of patient portals: what the literature tells us. Oakland, CA: California Health Care Foundation; 2011. URL: <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/M/PDF%20MeasuringImpactPatientPortals.pdf> [accessed 2013-03-26] [WebCite Cache ID 6FPRUiDVC]
63. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. *J Med Internet Res* 2013;15(2):e43 [FREE Full text] [doi: [10.2196/jmir.2243](https://doi.org/10.2196/jmir.2243)] [Medline: [23470453](https://pubmed.ncbi.nlm.nih.gov/23470453/)]
64. Bigne E, Ruiz C, Sanz S. The impact of internet user shopping patterns and demographics on consumer mobile buying behaviour. *Journal of Electronic Commerce Research* 2005;6(3):193-209.
65. Li H, Kuo C, Russell M. The impact of perceived channel utilities, shopping orientations, and demographics on the consumer's online buying behavior. *J Comput Mediat Commun* 1999;5(2):1-20. [doi: [10.1111/j.1083-6101.1999.tb00336.x](https://doi.org/10.1111/j.1083-6101.1999.tb00336.x)]
66. Liao Z, Cheung MT. Internet-based e-shopping and consumer attitudes: an empirical study. *Information & Management* 2001 Apr;38(5):299-306. [doi: [10.1016/S0378-7206\(00\)00072-0](https://doi.org/10.1016/S0378-7206(00)00072-0)]
67. 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/)]
68. Källander K, Tibenderana JK, Akpogheneta OJ, Strachan DL, Hill Z, ten Asbroek AH, et al. Mobile health (mHealth) approaches and lessons for increased performance and retention of community health workers in low- and middle-income countries: a review. *J Med Internet Res* 2013;15(1):e17 [FREE Full text] [doi: [10.2196/jmir.2130](https://doi.org/10.2196/jmir.2130)] [Medline: [23353680](https://pubmed.ncbi.nlm.nih.gov/23353680/)]
69. United States Census Bureau. 2012. US Census Data URL: <http://www.census.gov/> [accessed 2013-07-02] [WebCite Cache ID 6HoeZ5Nxv]
70. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. *J Appl Psychol* 2001 Feb;86(1):114-121. [Medline: [11302223](https://pubmed.ncbi.nlm.nih.gov/11302223/)]
71. Anderson JC, Gerbing DW. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin* 1988;103(3):411-423. [doi: [10.1037//0033-2909.103.3.411](https://doi.org/10.1037//0033-2909.103.3.411)]
72. Hair JF. *Multivariate Data Analysis with Readings*. Englewood Cliffs, NJ: Prentice Hall; 1995.
73. Segars A. Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research. *Omega* 1997 Feb;25(1):107-121. [doi: [10.1016/S0305-0483\(96\)00051-5](https://doi.org/10.1016/S0305-0483(96)00051-5)]
74. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 1981 Feb;18(1):39. [doi: [10.2307/3151312](https://doi.org/10.2307/3151312)]
75. Nunnally J, Bernstein IH. *Psychometric Theory*. New York: McGraw-Hill; 1994.
76. Aiken LS, West SG, Reno RR. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications; 1991.
77. MyFitnessPal.com. iTunes. 2013. Calorie Counter & Diet Tracker by MyFitnessPal URL: <https://itunes.apple.com/us/app/calorie-counter-diet-tracker/id341232718?mt=8&ign-mpt=uo%3D4> [accessed 2013-07-11] [WebCite Cache ID 6I2FTkmqP]
78. Azumio Inc. iTunes. 2013. Instant Heart Rate - Heart Rate Monitor URL: <https://itunes.apple.com/us/app/instant-heart-rate-heart-rate/id409625068?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2FX0yh2]
79. US Preventive Medicine, Inc. iTunes. 2013. Macaw URL: <https://itunes.apple.com/us/app/macaw/id480193633?mt=8> [accessed 2013-07-11] [WebCite Cache ID 6I2FaYwe6]
80. 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/)]
81. Houston TK, Allison JJ. Users of Internet health information: differences by health status. *J Med Internet Res* 2002;4(2):E7 [FREE Full text] [doi: [10.2196/jmir.4.2.e7](https://doi.org/10.2196/jmir.4.2.e7)] [Medline: [12554554](https://pubmed.ncbi.nlm.nih.gov/12554554/)]
82. Boulanger B, Kearney P, Ochoa J, Tsuei B, Sands F. Telemedicine: a solution to the followup of rural trauma patients? *J Am Coll Surg* 2001 Apr;192(4):447-452. [Medline: [11294400](https://pubmed.ncbi.nlm.nih.gov/11294400/)]

Abbreviations

AVE: average variance extracted

CFA: confirmatory factor analysis

CMP: compatibility
HLTH: perceived healthiness
IT: information technology
OLS: ordinary least squares
PIIT: personal innovativeness toward information technology
PIMS: personal innovativeness toward mobile services
SES: socioeconomic status
VULN: perceived vulnerability

Edited by G Eysenbach; submitted 30.03.13; peer-reviewed by X Zhang, T Chomutare; comments to author 23.04.13; revised version received 10.05.13; accepted 16.05.13; published 02.08.13.

Please cite as:

Rai A, Chen L, Pye J, Baird A

Understanding Determinants of Consumer Mobile Health Usage Intentions, Assimilation, and Channel Preferences

J Med Internet Res 2013;15(8):e149

URL: <http://www.jmir.org/2013/8/e149/>

doi: [10.2196/jmir.2635](https://doi.org/10.2196/jmir.2635)

PMID: [23912839](https://pubmed.ncbi.nlm.nih.gov/23912839/)

©Arun Rai, Liwei Chen, Jessica Pye, Aaron Baird. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 02.08.2013. 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-Based Strength-Balance Training to Motivate and Improve Adherence to Exercise in Independently Living Older People: A Phase II Preclinical Exploratory Trial

Patrícia Silveira^{1*}, PhD; Rolf van de Langenberg^{2*}, PhD; Eva van het Reve^{2*}, MSc; Florian Daniel^{1*}, PhD; Fabio Casati^{1*}, PhD; Eling D de Bruin^{2*}, PhD

¹University of Trento, Department of Information Engineering and Computer Science, Trento, Italy

²Institute of Human Movement Sciences and Sport, Department of Health Sciences and Technology, ETH Zürich, Zürich, Switzerland

* all authors contributed equally

Corresponding Author:

Patrícia Silveira, PhD

University of Trento

Department of Information Engineering and Computer Science

Via Sommarive, 14-Povo

Trento, 38123

Italy

Phone: 39 0461 28 20 93

Fax: 39 0461 28 20 93

Email: silveira@disi.unitn.it

Abstract

Background: Reaction time, coordination, and cognition performance typically diminish in older adults, which may lead to gait impairments, falls, and injuries. Regular strength–balance exercises are highly recommended to reduce this problem and to improve health, well-being, and independence in old age. However, many older people face a lack of motivation in addition to other strong barriers to exercise. We developed ActiveLifestyle, an information technology (IT)-based system for active and healthy aging aiming at improving balance and strength. ActiveLifestyle is a training app that runs on a tablet and assists, monitors, and motivates older people to follow personalized training plans autonomously at home.

Objective: The objectives were to (1) investigate which IT-mediated motivation strategies increase adherence to physical exercise training plans in older people, (2) assess the impact of ActiveLifestyle on physical activity behavior change, and (3) demonstrate the effectiveness of the ActiveLifestyle training to improve gait speed.

Methods: A total of 44 older adults followed personalized, 12-week strength and balance training plans. All participants performed the exercises autonomously at home. Questionnaires were used to assess the technological familiarity and stage of behavior change, as well as the effectiveness of the motivation instruments adopted by ActiveLifestyle. Adherence to the exercise plan was evaluated using performance data collected by the app and through information given by the participants during the study. Pretests and posttests were performed to evaluate gait speed of the participants before and after the study.

Results: Participants were 75 years (SD 6), predominantly female (64%), held a trade or professional diploma (54%), and their past profession was in a sitting position (43%). Of the 44 participants who enrolled, 33 (75%) completed the study. The app proved to assist and motivate independently living and healthy older adults to autonomously perform strength–balance exercises (median 6 on a 7-point Likert scale). Social motivation strategies proved more effective than individual strategies to stimulate the participants to comply with the training plan, as well as to change their behavior permanently toward a more physically active lifestyle. The exercises were effective to improve preferred and fast gait speed.

Conclusions: ActiveLifestyle assisted and motivated independently living and healthy older people to autonomously perform strength–balance exercises over 12 weeks and had low dropout rates. The social motivation strategies were more effective to stimulate the participants to comply with the training plan and remain on the intervention. The adoption of assistive technology devices for physical intervention tends to motivate and retain older people exercising for longer periods of time.

(*J Med Internet Res* 2013;15(8):e159) doi:[10.2196/jmir.2579](https://doi.org/10.2196/jmir.2579)

KEYWORDS

motivation; exercises; aged; tablet; mobility; delivery of health care

Introduction

Background

The primary goal of public health care is to increase the number of years of good health and maintain independence and quality of life as long as possible. Healthy aging is characterized by the avoidance of disease and disability, the maintenance of high physical and cognitive function, and sustained engagement in social and productive activities. These 3 components together define successful aging [1].

An important part of successful aging is maximization of physical performance. The ability to fully participate in productive and recreational activities of daily life may be affected when the capacity to easily perform common physical functions decreases [1]. Thus, health status is an important indicator of quality of life among older people [2,3]. It has been demonstrated that components of health-related fitness and functional performance or serious, chronic conditions and diseases that directly influence the components of fitness and performance are related to perceived health among middle-aged and older adults [3-5].

Regular physical activity or exercise substantially prevents the development and progression of most chronic degenerative diseases [6-8], is of benefit to frail and older persons, and is the only therapy found to simultaneously improve sarcopenia, physical function, cognitive performance, and mood in older adults [9]. For older people, a sedentary lifestyle also increases the risk of falls, whereas physically active older people have a reduced risk of falls with injuries [10-12]. An important marker for improvements in physical function that influences health and survival is gait speed [13]. In summary, to increase older adults' quality of life and fitness, we need to encourage them to become or stay physically active [14-15] and increase their fitness through training.

The objective of this research is to run a phase II study [16] with a tablet app called ActiveLifestyle [17], an app for the autonomous strength-balance physical training for independently living older adults. We aimed to investigate (1) which information technology (IT)-mediated motivation strategies increased adherence to physical exercise training plans in older people, (2) whether these strategies could induce physical activity behavior change, and (3) the effectiveness of ActiveLifestyle training to improve gait speed.

Related Work

Home environmental interventions to prevent functional decline seem to be effective [18] and are preferred by older people (ie, instead of leaving their houses to exercise) [19]. Interventions with integrated assistive technology devices have, in this context, the potential to further help in overcoming some of the barriers to start training [20] and maintaining physical independence for independently living older people [21]. Recently developed innovative ideas designed to alter clinical practice in sports were based on the development of tablet apps

for prevention, for instance [22]. Tablet and smartphone software apps specifically designed for health purposes are, in general, enthusiastically adopted as a means of delivering self-managed health interventions [23-25]. However, such tablet-based interventions are often plagued by high attrition rates and varying levels of user adherence [24,25]. Furthermore, the effectiveness of tablet-based health intervention approaches has not yet been demonstrated in older people.

From a pilot study, we knew it is feasible to use assistive technology devices in an older population with the aim of encouraging performance of physical exercise [20]. The short-duration pilot did not focus on aspects of physical functioning, but indicated that the app could be improved by explicitly considering additional motivational strategies. It is well known that motivation strategies affect adherence to health interventions [26]; however, only a few solutions explore different motivation techniques to stimulate regular physical activity [26-28]. Most of these solutions have the drawback that they do not specifically focus on older people. Albaina et al [29] presented a user-friendly software interface running on a small touchscreen display to motivate older adults to walk. The authors used a graphical representation of a flower to motivate and assist seniors to monitor their daily amount of steps collected by pedometers through this simple metaphor of their performance. To the best of our knowledge, there is not another software app dedicated to strength-balance training plans for older people.

Methods

ActiveLifestyle

ActiveLifestyle is a software app for active aging, aimed at assisting, monitoring, and motivating older people during autonomous home-based physical workouts [20,30,31]. The software takes usability aspects into account, to ensure that older users can use it independently and it adopts a set of motivation strategies to stimulate users to exercise regularly. A video of the app is available on YouTube [32], and the app can be downloaded from the Apple App Store [17].

Three levels of strength-balance training plans are supported in the app: beginner, intermediate, and expert. In all levels, the balance training should be done 5 days per week. Sessions are composed of 3 exercises, in which the trainees repeatedly (1-3 times) hold a certain position for several seconds (15-30 sec). Each level has different exercises, allowing progression as the person advances through the levels (eg, at the intermediate level the older person must perform the exercises while standing on a towel; at the expert level the exercise must be performed with the eyes closed). Strength training has 3 levels and should be done twice a week; starting with 6 warm-up exercises, then 10 strength and 2 stretching exercises. A minimum number of sets (1-3) and repetitions (12-30) are available for each exercise. Some exercises require the use of weights (2-6 kg). The required effort of the exercises increases according to the level (eg, the beginner level does not require weights; the intermediate level

requires ankle weights and performance in the sitting position; the expert level requires weights and exercises performed in standing position). The strength-balance training follows best practices recommendations and training principles (eg, it is progressive in nature) [33,34]. Figure 1 illustrates some exercises supported by ActiveLifestyle. All exercises are available on YouTube [35].

In addition to the actual physical training, ActiveLifestyle features a set of individual and social motivation instruments. In general, individual motivation strategies aim to convince someone to do something because it is inherently enjoyable for this person, independently of any social pressure. ActiveLifestyle specifically supports:

1. Conditioning through positive and negative reinforcement by immediately offering a reward/praise after an expected behavior to encourage the behavior and increase the probability that it happens again, or reprimands undesired behavior to decrease the probability of a reoccurrence of that behavior. Metaphors for reinforcement include a flower that grows whenever a session is completed (ie, positive reinforcement) and a gnome who takes care of the flower. The gnome's mood status varies according to the person's daily compliance to the plan (ie, positive and negative reinforcement; if the person performs the exercise, the gnome is happy, otherwise he is sad) (see panel a in Figure 2).
2. Goal setting by establishing specific, measurable, achievable, and time-targeted goals. The goal is anticipated by visually conveying the achievable maximum growth of the flower (see panel b in Figure 2).
3. Self-monitoring by allowing people to monitor themselves and to modify their attitudes and behaviors. Coloring the flower growth stages reflects progress toward the goal (see panel b in Figure 2).
4. Awareness by presenting the benefits of being physically active through written content on a bulletin board and by showing inspiring stories (eg, link to newspapers, videos, or websites) (Figure 3).

Social motivation strategies are built on social psychology. An individual's social network (other trainees) may act as source of motivation. ActiveLifestyle uses:

1. Comparison by allowing a person to compare similarities and differences between 2 or more parties. People tend to keep equality in their relationships. Whenever a person completes a workout session, an automatic message is posted on a bulletin board informing the training community (ie, other users following the same training plans) about the complete session. The message also carries the status of the individual's flower.
2. External monitoring by allowing 1 party to monitor the performance of another party. ActiveLifestyle enables health care experts to access data on performance and compliance with the training plan. The older users have access to their own flower and to that of their training partners, enabling monitoring progress of peers.

3. Emotional support by encouraging exchange of written messages between trainees and experts to motivate and assist. ActiveLifestyle uses a bulletin board and an "inBox." The first is a public channel where all members of the training community have access. The second is a bidirectional private channel for contact with professionals capable of giving advice and feedback on trainings.
4. Collaboration by offering a collaborative activity designed as a game, in which to progress in the game, a group of trainees must jointly be compliant with the training plan. The To the Top game is a trekking trail with 24 predefined points (2/week). The aim of the game is to climb a mountaintop, as a group of successful trainings ends. Compliance with the training plan is evaluated twice weekly on group level. A total of 65% or more members of a group have to perform the scheduled workout to be awarded a new flag on the trail (representing progress toward the mountaintop). Each flag uncovers a story with trivia about the Matterhorn and what is needed to conquer the mountain as a parable explaining the benefits associated with being physically active (Figure 4).

ActiveLifestyle comes in 2 versions. The individual version contains only the individual motivations strategies. The social version supports individual and social motivation strategies, and a virtual training plan community and communication features. In addition to the motivation strategies, ActiveLifestyle supports 6 main features accessible through its menu:

1. The What's Next? option invites the users to start the performance of due workout sessions.
2. The weekly exercises option shows the scheduled strength-balance sessions organized per week.
3. The progress option shows the users' progress through the conditioning, goal setting, and self-monitoring strategies previously mentioned in both versions. The social version also supports the collaboration strategy through the To the Top game.
4. The bulletin board allows the users to receive written messages, which may include links for websites and YouTube videos. Three types of messages are supported: (1) workout session completed messages (in green) to inform the participant(s) about the conclusion of a scheduled session of exercises; (2) ActiveLifestyle tips messages (in pink) to support the awareness motivation strategy illustrated in Figure 3; and (3) public messages (in white) written by the training members. Only the social version supports the third type of message and has the ability to send messages to the entire training plan community.
5. The friends option lists the members of the training plan community (ie, older users and experts). Only the social version supports this feature.
6. The inBox option allows users to exchange private text messages with their list of friends.

All the previously mentioned features and motivation strategies can be inspected at the Life Participation Project website [31].

Figure 1. Exercise examples.

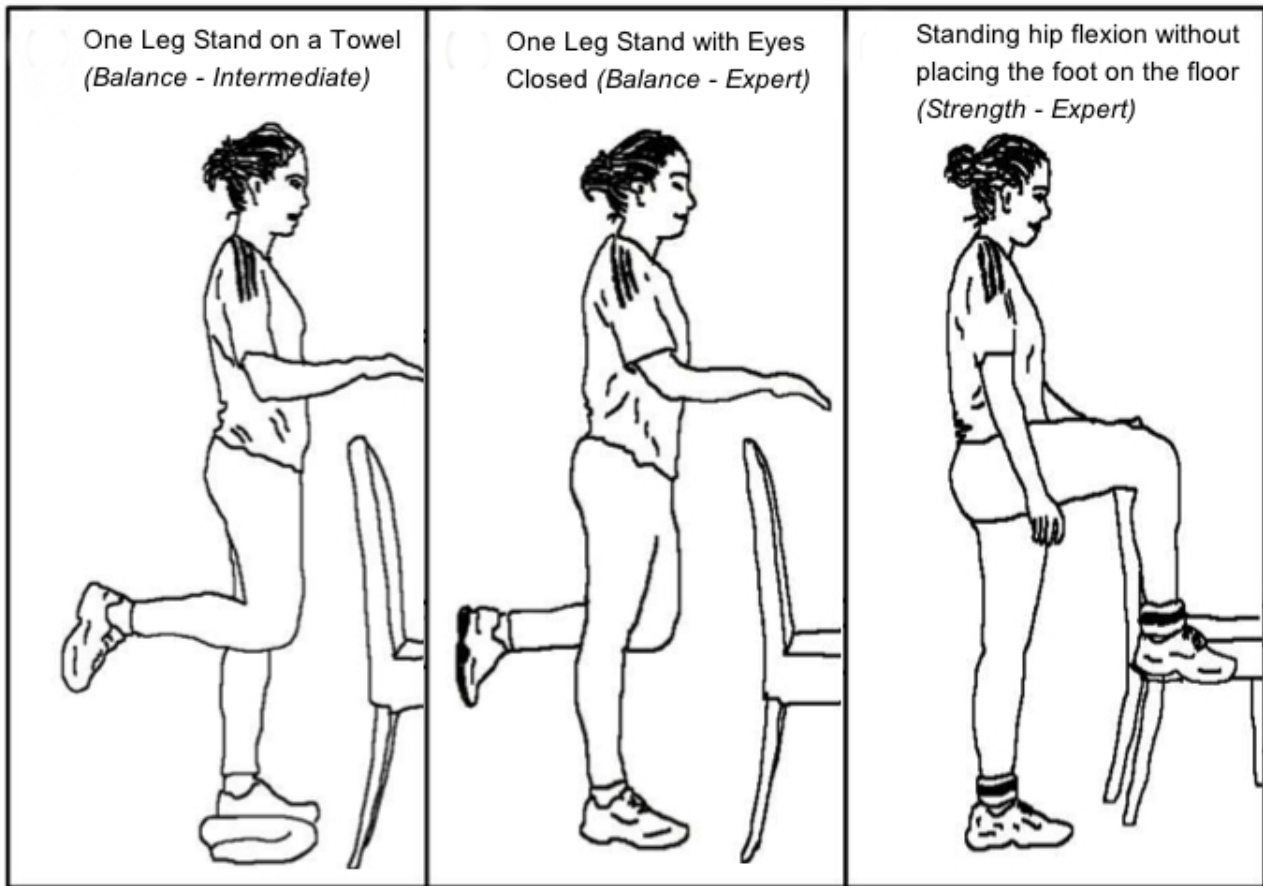


Figure 2. Metaphors within the app to motivate older people through conditioning, goal setting, and self-monitoring.



Figure 3. ActiveLifestyle tips to improve awareness about the benefits of being physically active.

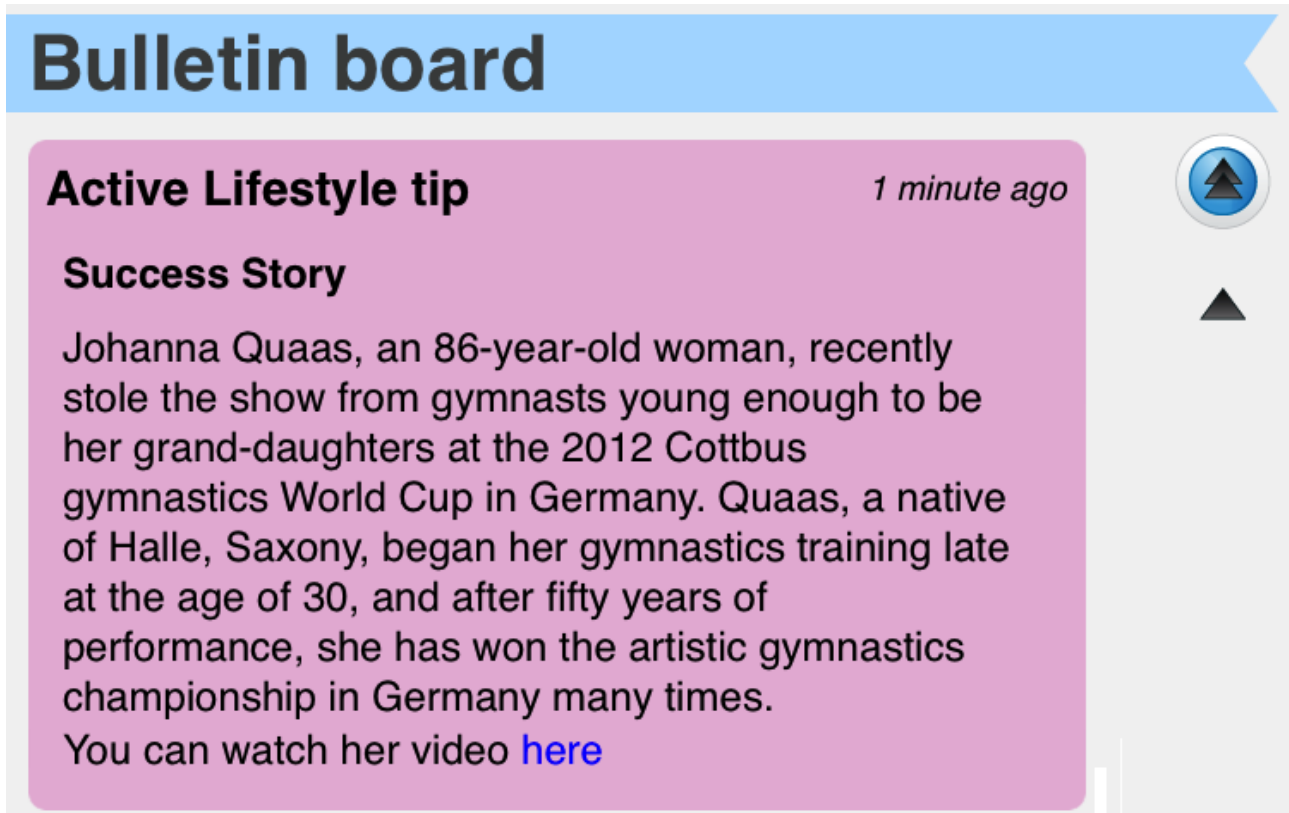
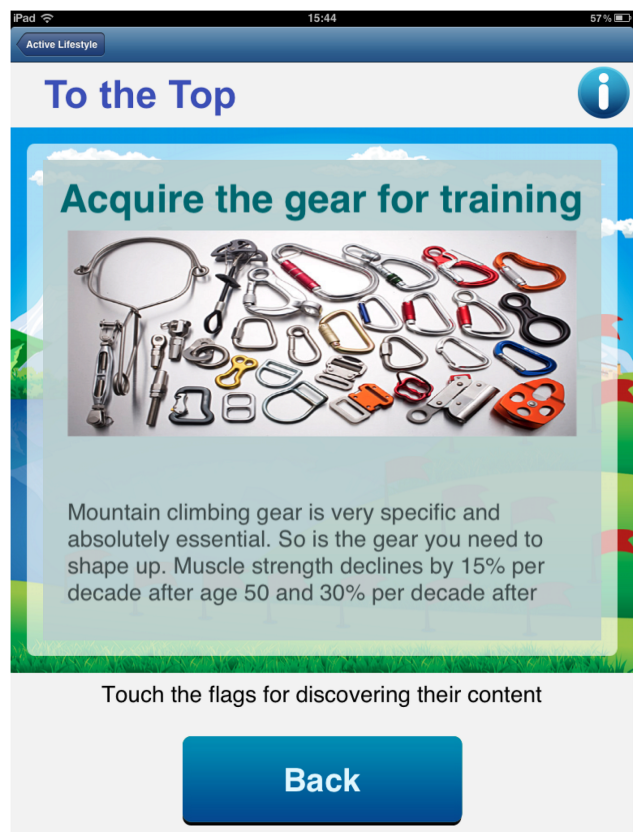
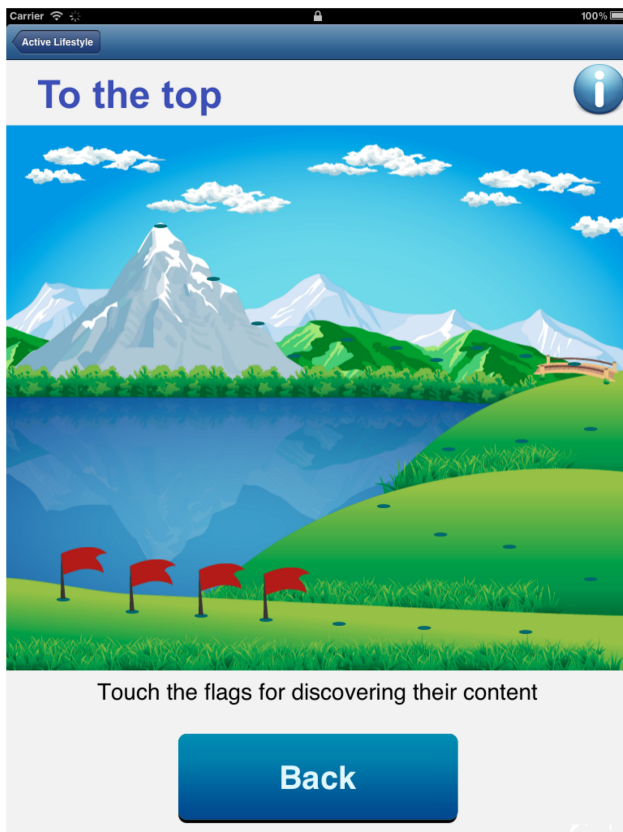


Figure 4. The To the Top collaboration game within the app.



Eligibility Criteria

Participants were older adults aged 65 years or older; living independently; able to walk independently with or without walking aids; able to follow instructions spoken in German, English, or Italian; and with no severe illness, cognitive impairment, progressive neurological disease, stroke, severe cardiac failure, or high blood pressure. Ethical approval for the study was obtained from the Eidgenössische Technische Hochschule (ETH) Ethics Committee (EK 2011-N-64).

Setting

Participants were recruited by convenience sampling from 2 institutions for older people and 1 organization responsible for coordinating and providing at-home nursing care for seniors. The Senioren Begegnungszentrum Baumgärtlihof, a day center dedicated to delivering services and information related to the older population (Horgen, Switzerland), advised potential participants through its mailing list and by notes in the local newspaper (Multimedia Appendix 1). The Alterswohnungen Turm-Matt, a cooperative offering housing and daily living facilities to older people (Wollerau, Switzerland), informed and advised potential participants in person or by phone and distributed flyers to advertise the study. The Fachstelle für präventive Beratung Spitex-Zürich, a home-care nursing organization (Spitex-Zürich), promoted the study by sending letters and specifically inviting patients in need of better physical performance. Spitex-Zürich nurses selected potential participants based on the eligibility criteria.

Intervention

To investigate the effects of different motivation strategies, a pretest/posttest preclinical trial was performed. For convenience, the ActiveLifestyle groups were composed of (1) an individual group that followed training using the individual version of ActiveLifestyle; (2) a social group that followed training using the social version of the app, (3) a control group that followed exercises with printed information without additional motivation strategy. The individual and social groups were randomly composed of participants recruited from Baumgärtlihof and Spitex-Zürich, whereas the participants in the control group were recruited from Turm-Matt because of time and resource constraints (eg, lack of research team members, the control group was not randomized with the other participants). Figure 5 shows the recruitment process and the flow of participants through the study. Videos of some parts of the interventions can be watched on YouTube [35].

The development of our intervention follows a framework for the design and evaluation of complex interventions [36] and should at this stage be considered as a preclinical exploratory trial. For this reason, we did not use a pure randomized, controlled research design; therefore, we did not register this study as a clinical trial.

Outcome Measures

Adherence and Attrition

Adherence was computed by ActiveLifestyle during the intervention and stored in a central database. The control group adherence was assessed with paper-based training logs. To

calculate adherence, the total number of workout sessions for each participant was divided by 81, which was the total number of possible training sessions for the 12-week period (because of technical issues, the training was suspended for 3 days and the trainees were aware of the 81 training sessions in advance). The adherence of participants who dropped out was calculated by dividing the number of workout sessions attended up to the point of dropout from the study by 81 [37]. Values were compared between groups and with median rates in community-based fall prevention interventions [38]. For attrition, we measured the number of participants retained and lost at the final follow-up.

Gait Speed

The effect of the training on physical performance was assessed by measuring preferred and fast walking speed [39] with the GAITRite walkway, a valid and reliable tool for measuring gait in older people [40-42].

Motivation Instruments

The effectiveness of the motivation instruments built into the system was assessed based on the participants' feedback, collected with a 7-point Likert scale self-reported questionnaire at the end of the intervention (Multimedia Appendixes 2 and 3), and on the performance (adherence, attrition, and gait speed) comparison among the 3 groups of participants.

Change of Behavior

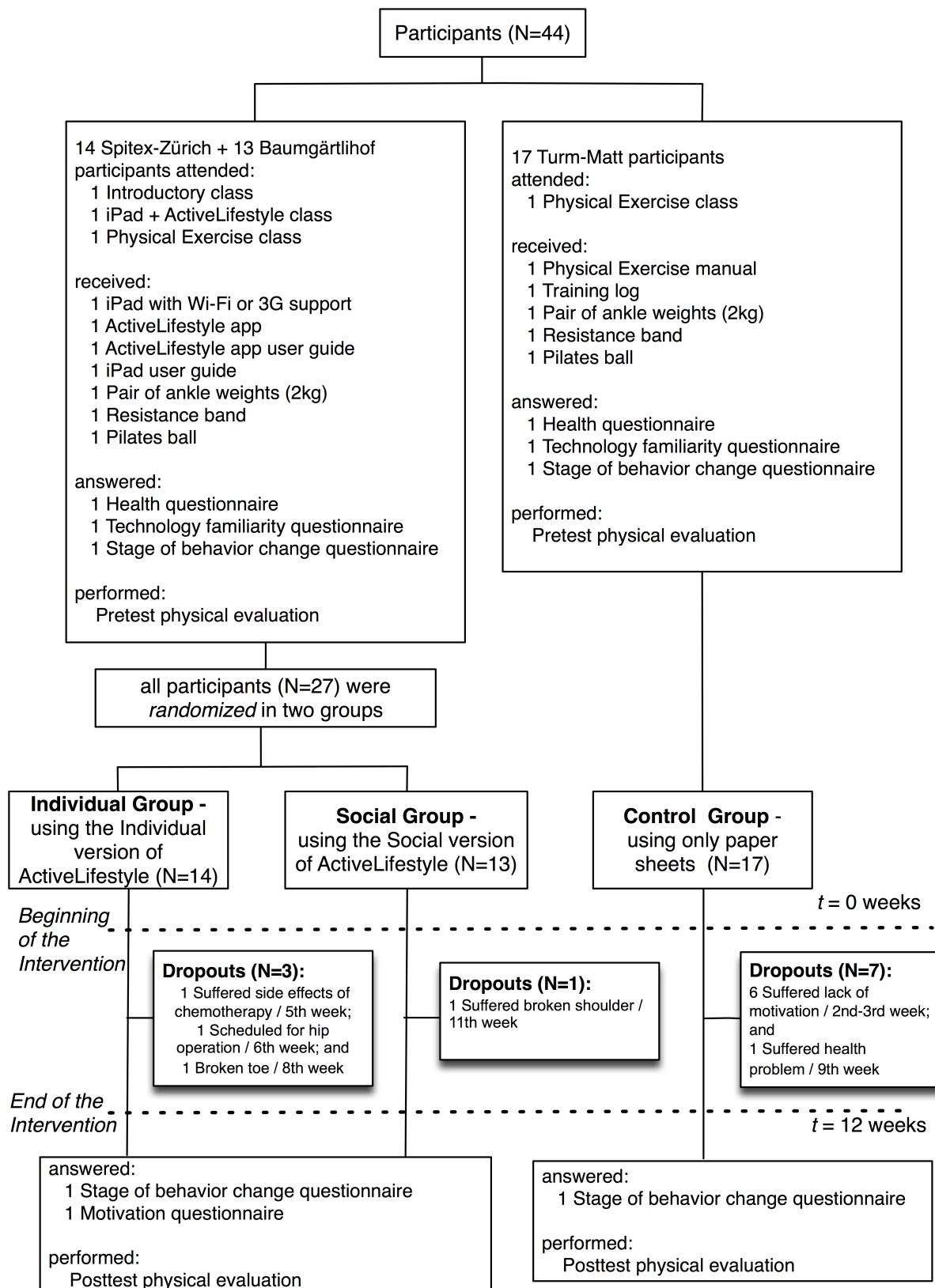
The level of exercise adoption was evaluated according to the Transtheoretical Model (TTM) [43], which describes how people modify or acquire behavior. A self-reported TTM questionnaire (Multimedia Appendix 4) was applied before and after the training period. Participants were classified into 4 groups: contemplation (eg, thinking about physical behavior change), preparation (eg, already somewhat physically active), action (eg, doing enough physical activity), and maintenance (eg, making physical activity a habit).

Statistical Analyses

Analysis of variance (ANOVA) was used to test for differences in adherence to the training program between groups, as well as gait speed over time and between groups. Significant main effects were followed up by post hoc *t* tests with correction for multiple comparisons. Between-group differences in attrition were analyzed using a chi-square (χ^2) test. Questionnaires on enjoyment, motivation, and change of behavior were analyzed using Kruskal-Wallis ANOVA and Wilcoxon signed rank tests (*W*). In all analyses, the level of significance was set at $P \leq .05$.

For effect size, we used η^2 in all ANOVA analyses, Cohen's *d* for all post hoc analyses, mean square contingency coefficient (ϕ) for chi-square tests, and Pearson *r* (*r*) for Kruskal-Wallis ANOVA and Wilcoxon signed rank tests. The *r* is calculated as $r = Z/\sqrt{N}$, in which *Z* is the standardized difference and *N* is the total number of samples. Suggested norms for interpreting η^2 are 0.01=small, 0.06=moderate, and 0.14=large effect [44]. For small, moderate, and large effects, these norms are 0.2, 0.5, and 0.8, respectively, for Cohen's *d* and 0.1, 0.3, and 0.5, respectively, for both ϕ and *r* [44]. All tests were conducted using SPSS Version 21.0 (IBM Corp, Armonk, NY, USA).

Figure 5. Flowchart of participants.



Results

Demographics

Detailed information about the participants' demographics, based on the Health and Technology Familiarity self-reported questionnaires (Multimedia Appendixes 5 and 6), is summarized in Table 1.

Adherence and Attrition

Table 2 presents the adherence to ActiveLifestyle strength–balance training plans. Adherence across training plans differed significantly between groups ($F_{2,41}=4.8$, $P=.01$, $\eta^2=0.19$). Post hoc t tests with Benjamini–Hochberg correction revealed a large and significant difference between the social group (mean 81.9%, SD 1.6%) and the control group (mean

48.1%, SD 41.5%; $t_{19,2}=3.1$, $P=.02$, $d=0.91$). The difference between the individual group (mean 71.1%, SD 25.2%) and the control group was moderate to large ($t_{26,9}=1.9$, $P=.10$, $d=0.63$). The difference between the individual and social groups was moderate yet nonsignificant ($t_{18,6}=1.4$, $P=.19$, $d=0.50$).

Thirty-three older adults completed the 12 weeks of training, resulting in a 25% attrition rate in total, 21% in the individual group (3/14), 8% in the social group (1/13), and 41% in the control group (7/17). Figure 6 illustrates the number of remaining participants in each group per week after enrollment. More details about the dropout reasons are reported in Figure 5. A chi-square test revealed that attrition rate was higher in the control group (41.2%) than in the combined ActiveLifestyle groups (14.2%; $\chi^2_1=3.9$, $P=.05$, $\phi=0.30$).

Table 1. Participants' demographics (N=44).

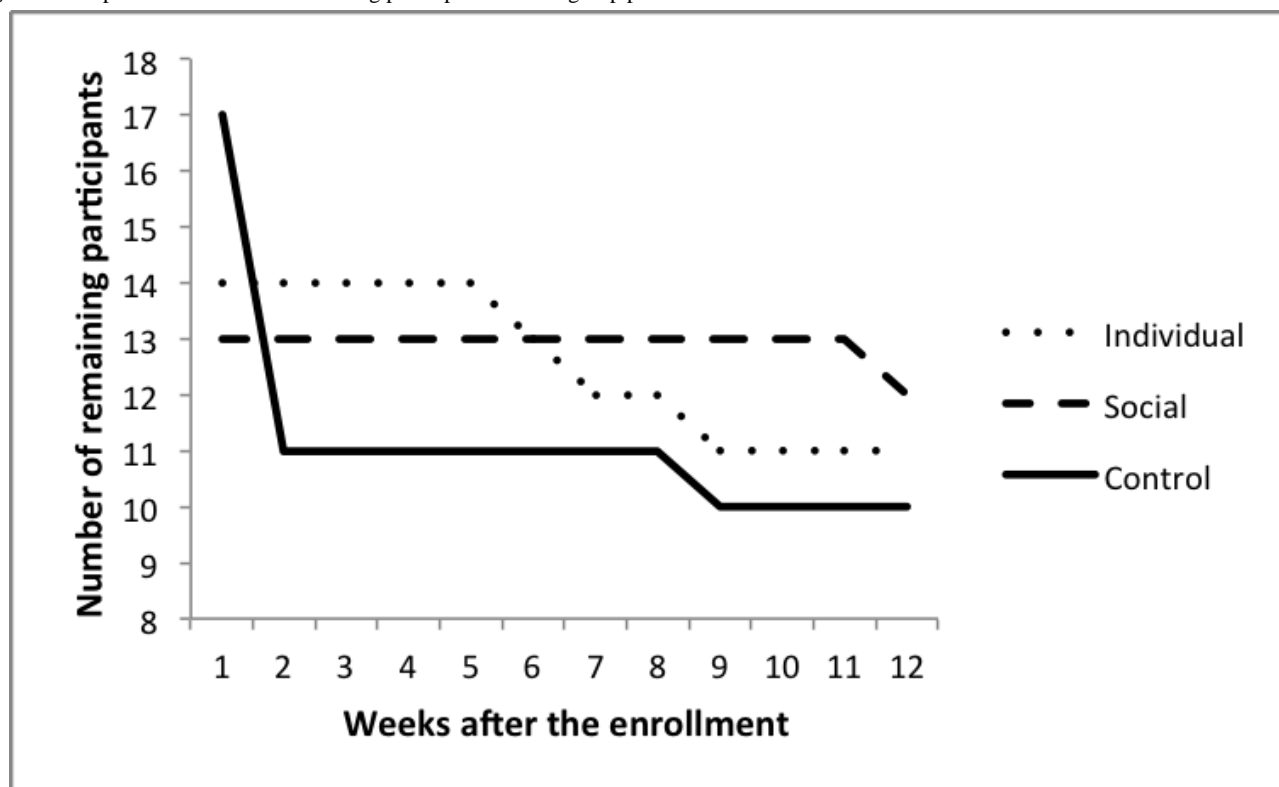
Characteristic	Individual (n=14)	Social (n=13)	Control (n=17)
Female gender, n (%)	10 (71)	8 (62)	10 (59)
Age (years), mean (SD)	74 (5)	75 (6)	76 (15)
Hold trades or professional diploma, n (%)	7 (50)	7 (54)	10 (59)
In a sitting position past profession, n (%)	7 (57)	6 (46)	6 (35)
Health questions, n (%)			
Estimated good health	5 (36)	8 (61)	8 (47)
Estimated average balance	7 (50)	5 (38)	9 (53)
Feel pain but not every day	9 (64)	7 (54)	7 (41)
Flexibility questions, n (%)			
Fell in the past 6 months ^a	2 (14)	5 (38)	4 (23)
Walk at least twice a week	5 (36)	8 (61)	9 (53)
Practiced some sport in the past	10 (71)	8 (61)	5 (29)
Never practiced strength exercises	11 (79)	7 (54)	14 (82)
Technology familiarity, n (%)			
Frequently use automated teller machines	7 (50)	9 (69)	7 (41)
Frequently use cellphones	7 (50)	10 (77)	6 (35)
Frequently use digital photography	8 (57)	4 (31)	4 (23)
Don't use Global Positioning System devices	7 (50)	8 (61)	6 (35)
Don't use automatic kiosks	9 (64)	6 (46)	12 (71)
Don't know what an e-book is	7 (50)	5 (38)	11 (65)
Use a computer	12 (86)	10 (77)	8 (47)
Between 1-5 hours per week	6 (43)	4 (31)	3 (18)
Use the Internet	12 (86)	9 (69)	5 (29)
Between 1-5 hours per week	7 (50)	6 (46)	2 (12)

^aA fall was defined as unintentionally coming to the ground or some lower level, excluding the consequence of sustaining a violent blow, loss of consciousness, or sudden onset of paralysis, such as during a stroke or epileptic seizure [45].

Table 2. Adherence to ActiveLifestyle strength–balance training plans.

Training plan	Individual group			Social group			Control group		
	Visited	Planned	%	Visited	Planned	%	Visited	Planned	%
Balance training plan	547	812	67	549	754	73	451	986	46
Strength training plan	221	322	69	217	299	73	291	391	74
Across training plans	768	1134	68	766	1053	73	742	1377	54

Figure 6. Graph of the number of remaining participants in each group per week.



Gait Speed

Table 3 shows participants’ preferred and fast gait speed during the pretest and posttest evaluations. With respect to preferred gait speed, the 3 groups were similar. We used 2 mixed 2-way ANOVA’s (1 for preferred and 1 for fast gait) with within-subject factor pre–post (2 levels) and between-subject factor group (3 levels). For preferred gait speed, there was a significant difference between pretest and posttest ($F_{1,29}=29.5$, $P<.001$, $\eta^2=0.50$). Participants walked significantly faster in the posttest (1.276 m/s) than they did in the pretest (1.142 m/s). There was no significant main effect of group ($P=.07$) and no significant interaction effect ($P=.65$), suggesting that preferred gait speeds and their improvements were similar in all groups.

The results for fast gait speed were similar to those for preferred gait speed. Again, there was a large difference between pretest and posttest: Participants walked significantly faster in the posttest (1.72 m/s) than in the pretest (1.56 m/s; $F_{1,29}=20.1$, $P<.001$, $\eta^2=0.41$). The main effect of group was significant also ($F_{2,29}=5.3$, $P=.01$, $\eta^2=0.27$). Post hoc tests revealed that the individual group (1.89 m/s) was significantly faster than the control group (1.45 m/s; $t_{19}=3.94$, $P=.003$, $d=1.31$), and faster

than the social group (1.58 m/s; $t_{20}=2.05$, $P=.08$, $d=.89$), though not statistically significant. The individual group, by chance, was the fastest from the beginning. Fast gait speed was not significantly different between the control group and the social group ($P=.39$).

Motivation Instruments

Detailed information about the effectiveness of ActiveLifestyle’s motivation instruments and user-intention aspects are summarized in Table 4. The questionnaires used to collect the content of the table are available in Multimedia Appendixes 2 and 3.

Most participants affirmed that ActiveLifestyle facilitates the autonomous performance of balance–strength exercises. This was confirmed by a high intention to use the app again or to recommend it to friends or family members. The individual group was unanimous in the evaluation of these 2 user-intention aspects, whereas the social group presented high values but not with unanimity. In general, the participants of both groups did not feel motivated to perform physical exercises before the study.

All the participants thought it was fun to perform the strength and balance exercises. Few participants (<25%) felt frustrated, worried, or nervous during the study. More than half of the participants, 54% from the individual group and 67% from social group, will miss ActiveLifestyle.

The individual motivation strategies seemed to be more effective on the individual group level than on the social group level. Most of the individual group felt motivated by the goal-setting and self-monitoring strategies (91%), both represented by the progress bar metaphor (panel b in Figure 2), as well as for being aware of the benefits of being physically active-aware (82%). Conditioning through positive and negative reinforcement also motivated the participants. In all, 64% felt motivated when they saw the plant growing, whereas 55% felt motivated by the mood status of the gnome.

The most effective motivating strategies for the social group were conditioning through positive social inclusion and external monitoring (all 83%). After that, the social group felt motivated through the awareness of the benefits of physically activity (82%), emotional support (75%), the monitoring of their progress toward the plan (goal setting and self-monitoring) (67%), participation in the collaboration game (58%), positive and negative reinforcement (conditioning) (50%), and the comparison of their performance with other training participants on the bulletin board (42%).

Most participants in the individual group (64%) expressed that they would feel more motivated if they could use the social version of ActiveLifestyle, but the reverse was not true. Only a few participants in the social group expected to be less motivated using the individual version of the app (8%).

Mann–Whitney *U* tests comparing the Likert scores for all questions presented in Table 4 did not detect any significant differences between the groups.

Change of Behavior

Table 5 shows the stage of behavior change of the participants at the beginning ($t_0=0$ weeks) and at the end ($t_1=12$ weeks) of the intervention.

Wilcoxon signed rank tests comparing pretest and posttest behavioral scores in each group revealed a trend—with a large effect size—in the social group ($W=1.79$, $P=.07$, $r=0.52$). Hence, the social group tended to change their behavior toward integration of ActiveLifestyle into their daily routine. No behavioral changes were detected in the control group ($P=.28$) or the individual group ($P=.50$). Although this suggests between group differences with respect to behavioral change, no such differences could be shown statistically; a Kruskal–Wallis ANOVA directly comparing change of behavior between the 3 groups was nonsignificant ($P=.75$).

Table 3. Participants' gait speed during the pretests and posttests.

Condition	Pretest mean (SD)	Posttest mean (SD)
Individual group		
Preferred speed (m/s)	1.26 (0.18)	1.42 (0.21)
Fast speed (m/s)	1.80 (0.27)	1.98 (0.31)
Social group		
Preferred speed (m/s)	1.10 (0.25)	1.24 (0.31)
Fast speed (m/s)	1.50 (0.35)	1.66 (0.50)
Control group		
Preferred speed (m/s)	1.07 (0.19)	1.17 (0.22)
Fast speed (m/s)	1.39 (0.22)	1.51 (0.27)

Table 4. Outcome data expressed by the participants on a 7-point Likert scale (range 1-7; 1=completely disagree to 7=completely agree) at the end of the intervention period.

Evaluation statements	Individual (n=14)		Social (n=13)	
	Median (range)	% Agreed	Median (range)	% Agreed
Statement				
ActiveLifestyle facilitates the performance of autonomous strength–balance exercises at home	7 (6-7)	100	7 (4-7)	92
Use intention				
I would use the app again	6 (5-7)	100	6 (4-7)	83
I would recommend the app to my friends and family	6 (6-7)	100	6 (3-7)	67
Enjoyment				
It was fun to carry out the strength and balance exercises	6 (6-7)	100	6 (5-7)	100
I felt frustrated during the study	2 (1-5)	9	2 (1-6)	8
I felt worried during the study	2 (1-6)	18	2 (1-7)	25
I felt nervous during the study	1 (1-6)	9	1 (1-4)	0
I will miss the exercises and the ActiveLifestyle app	5 (2-7)	54	6 (3-7)	67
Motivation				
I usually do not feel motivated to perform physical exercises, ActiveLifestyle helped me	6 (1-7)	54	6 (2-7)	83
Individual motivation instruments				
I felt motivated when I saw my performance on the progress bar (goal setting and self-monitoring)	6 (4-7)	91	6 (1-7)	67
I felt motivated by being aware about the benefits of being physically active (awareness)	6 (3-7)	82	6 (3-7)	82
I felt motivated when I saw the plant growing due to my performance (conditioning)	6 (4-7)	64	6 (1-7)	83
I felt motivated when I saw the emotional status of the gnome (conditioning)	5 (2-7)	55	4 (1-6)	50
I would feel more motivated using the social version of ActiveLifestyle, in which I could interact with other training partners	5 (1-7)	64	—	—
Social motivation instruments				
I felt motivated for being part of a training group and knowing that other people did the same exercises	—	—	6 (2-7)	83
I felt motivated to perform the plan because I knew I was being monitored (external monitoring)	—	—	6 (2-7)	83
I felt motivated for being emotionally supported by the other training partners and by the ActiveLifestyle experts (emotional support)	—	—	6 (2-7)	75
I felt motivated with the collaboration activity to reach the top of the mountain (collaboration)	—	—	6 (3-7)	58
I usually compared my flower with others on the bulletin board (comparison)	—	—	4 (1-6)	42
I would feel more motivated using the individual version of ActiveLifestyle, which does not require interaction with other training partners	—	—	4 (1-6)	8

Table 5. Stage of behavior change of the participants according to the Transtheoretical Model (TTM).

Stage of behavior change	t ₀ = 0 weeks	t ₁ = 12 weeks
Individual group		
Contemplation	3	0
Preparation	1	4
Action	1	1
Maintenance	6	6
Social group		
Contemplation	5	3
Preparation	2	0
Action	0	0
Maintenance	5	9
Control group		
Contemplation	5	3
Preparation	1	2
Action	0	0
Maintenance	4	5

Discussion

Principal Findings

The aim of this study was to investigate (1) which IT-mediated motivation strategies increase adherence to physical exercise training plans in older people, (2) whether the ActiveLifestyle app induces physical activity behavior change, and (3) the effectiveness of the ActiveLifestyle training to improve gait speed. The main focus was to evaluate the ability to retain older people in the exercise program. Based on findings from a systematic review [39], we could expect a 10% attrition rate and 50% adherence rate for the individually targeted exercise training. Although the control group showed 41% attrition (primarily because of lack of exercise motivation), both tablet-based training groups showed far lower values, 21% and 8% for the individual and social ActiveLifestyle groups, respectively. These last 2 numbers also contain the effect of morbidities not related to the motivation to train (ie, unexpected health problems). Especially in the control group participants, the lack of motivation for continuous training was high. The degree of engagement with the intervention was more than 68% for the individual group and 73% for the social group, both using the ActiveLifestyle app, and 54% for the control group. Compared with median rates for attrition (10%) and adherence (50%) in fall prevention interventions in community settings, we achieved better or similar rates for the tablet-based training groups. From previous research [46], we know that the intention to undertake strength-balance training in older people is closely related to all elements of coping appraisal. Elements of coping appraisal include the belief that strength-balance training has multiple benefits, a positive social identity, and the feeling that family, friends, and doctors would approve of taking part in such training [46]. It can be hypothesized that ActiveLifestyle is effective in influencing attrition and adherence because it explicitly supports individual and social motivation instruments.

The reason to use a tablet solution is related to the numerous potential advantages attributed to such a tool (eg, tablets are relatively robust, and using fingers instead of a mouse or a touch pad make them much more intuitive and easy to use compared with smartphones, notebooks, and desktops). A tablet-based intervention, such as ActiveLifestyle, constitutes a powerful tool to provide feedback about performance and motivation to endure practice because of social inclusion. Interventions that use frequent, nonfrequent, or direct remote feedback are to be favored versus treatments without feedback, because the former seem to be more effective than the latter and they are equally effective as supervised exercise interventions [47]. The second most-mentioned barrier to physical exercise for subjectively insufficiently active older adults is lack of company. Direct remote contact seems to be a good alternative to supervised on-site exercising [47]. Such feedback can easily be adapted to the individual participant's baseline motor performance and progressively augmented with task difficulty. ActiveLifestyle has been demonstrated to have the potential to engage people who otherwise would lack interest to participate in a physical exercise regimen. Especially in the older population, it is difficult to maintain high adherence to training programs [48]. The participants of the present study allocated to the tablet groups showed good compliance rates. The losses related to low exercise compliance (n=6) in the control training group were caused by a lack of motivation. The reasons for discontinuation of training in the tablet groups were not because of rejection of the app; they were because of health problems. In a future phase III trial, the follow-up period for the assessment of adherence and attrition should preferably be extended to 12 months to enable the comparability of this future study with reference values of previous physical interventions [39]. Although the result of a 12-week intervention, our findings are encouraging and indicate the effectiveness of a tablet-based

training approach in older people. This encourages further exploration of this training approach in seniors.

Analyzing the participants' answers to the motivation instruments of ActiveLifestyle, most of the individual participants (64%) would feel more motivated using the social version of the app, whereas the opposite is not true (8%) (both tablet groups were aware of the different versions of ActiveLifestyle). Regarding the physical activity habits, the training group using the social version of ActiveLifestyle was the only group showing a tendency to change behavior. At the end of the intervention, 50% of the social group participants changed their behavior according to the TTM. At the beginning, these participants were at the contemplation or preparation stages (thinking about or already being somewhat physically active), and they were classified as being on the maintenance stage (making physical activity a habit) by the end. However, a further longitudinal study with a larger sample, including evaluation after the end of the intervention, is required to be able to ascertain change of physical behavior.

Gait speed is a clinically relevant indicator of functional status associated with important geriatric health outcomes (ie, impact health care activities have on people) [49]. Slowing down has been recognized as an indicator of failing health and vulnerable old age [50]. Some researchers hypothesize that gait speed may act as a vital sign, giving indications of the health status of older people. Mortality, for example, is substantially reduced when gait speed is improved through interventions [51]. Large epidemiological studies reveal that a 0.1 m/s faster walking speed is related to a 12% decrease in mortality [13]. In this respect, it is encouraging that all older people in our training groups that adhered to their training plan, independently of their group allocation, showed an increase in both preferred and fast walking speed.

In addition to the high level of adherence caused by the social motivation instruments, the training community created by the study served to improve the connectedness of the participants, which may help people to garner social support for making physical changes in their daily lives [52]. Two women who did not know one another started to perform the exercises together to check if they were following the correct posture. Some participants contacted other training partners using the app or via email or phone when they faced problems. The same support was also requested from our team of experts, who frequently (especially at the beginning of the study) received phone calls because of technical problems or doubts about the exercises.

As learned in our previous study [20], some of the participants felt proud of being able to use new technology. One of our oldest participants (83 years) installed Skype to call his daughter living in Central America. He confessed that his daughter was very surprised. In the beginning, 1 woman was afraid of not being able to correctly operate the tablet because she had never used a computer before. After the study, she bought a tablet on her own to play with her grandchildren and installed Wi-Fi at home to be more connected with them. Another woman expressed a similar concern at the beginning of the training, but finished

the study with a new tablet and a Gmail account: "I'm proud to be in possession of the iPad and to be able to write to my friends. The whole matter was a change for me."

Limitations

The study has some limitations. One of them is the rather small sample size. The study reveals first estimates for gait speed measures and stages of behavior change and warrants further research in larger populations. However, the purpose of preclinical exploratory trials is to provide preliminary evidence on the clinical efficacy of an intervention [16,36]. When evaluating the validity of a study, it is important to consider both the clinical and statistical significance of the findings [53]. Studies that claim clinical relevance may lack sufficient statistical significance to make meaningful statements or, conversely, may lack practicality despite showing a statistically significant difference in treatment options. Researchers and clinicians should not focus on small *P* values alone to decide whether a treatment is clinically useful; it is necessary to also consider the magnitude(s) of treatment differences and the power of the study [53]. Encouraging in this context is the observation that most of the between-groups comparisons for adherence show medium or medium-to-high magnitude(s) of treatment differences in favor of the tablet groups. The relationship between tablet-based physical training research and its effect on adherence and fitness in older individuals requires further exploration. Another limitation of this study is related to the research design used. The different recruitment methods and the lack of initial randomization and blinding may have introduced a selection bias that questions the validity of the adherence/motivation findings. Analogous studies with similar or frailer populations and the use of a true randomized controlled research design should be performed to substantiate or refute our findings.

The participants of this study can be classified as normal walkers with a preferred gait speed between 1.0 and 1.4 m/s. Future studies with community dwelling populations that exhibit mildly abnormal (0.6-1.0 m/s) or seriously abnormal gait speed (<0.6 m/s) [50] should be performed to investigate whether similar or even better results in physical performance variables can be obtained.

Conclusion

The finding of this study supports the notion that it is advantageous to combine physical training with specifically targeted IT motivation instruments that offer the possibility to socialize in a group in clinical practice. The combination seems to have a positive influence on older adults' training adherence in comparison to more traditional exercise. ActiveLifestyle proved to assist and motivate independently living and healthy older adults to autonomously perform strength-balance exercises. The social motivation strategies seemed to be more effective to stimulate the participants to comply with the training plan and remain on the intervention. The adoption of assistive technology devices for physical intervention tends to motivate and retain older people exercising for longer periods of time.

Acknowledgments

This work was carried out with the support of Maria Messmer-Capaul and her team at the Fachstelle für präventive Beratung Spitex-Zürich, of Claudia Nüesch at the Senioren Begegnungszentrum Baumgärtlihof, and Data Quest, who provided tablets for the study. We would like to thank all the participants for their great enthusiasm and support. This work is partially sponsored by the European Institute of Innovation & Technology (EIT) project Virtual Social Gym and the EU FP7 SME Capacities program BPM4People.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Das iPad sagt Knie beugen: Zürichsee-Zeitung.

[[PDF File \(Adobe PDF File\), 209KB - jmir_v15i8e159_app1.pdf](#)]

Multimedia Appendix 2

Individual feasibility questionnaire.

[[PDF File \(Adobe PDF File\), 65KB - jmir_v15i8e159_app2.pdf](#)]

Multimedia Appendix 3

Social feasibility questionnaire.

[[PDF File \(Adobe PDF File\), 62KB - jmir_v15i8e159_app3.pdf](#)]

Multimedia Appendix 4

Transtheoretical-model questionnaire.

[[PDF File \(Adobe PDF File\), 49KB - jmir_v15i8e159_app4.pdf](#)]

Multimedia Appendix 5

Technology familiarity questionnaire.

[[PDF File \(Adobe PDF File\), 56KB - jmir_v15i8e159_app5.pdf](#)]

Multimedia Appendix 6

Health questionnaire.

[[PDF File \(Adobe PDF File\), 57KB - jmir_v15i8e159_app6.pdf](#)]

References

1. Rowe JW, Kahn RL. Successful aging. *The Gerontologist* 1997;37(4):433-440. [doi: [10.1093/geront/37.4.433](https://doi.org/10.1093/geront/37.4.433)]
2. Spirduso WW, Cronin DL. Exercise dose-response effects on quality of life and independent living in older adults. *Med Sci Sports Exerc* 2001 Jun;33(6 Suppl):598-608. [Medline: [11427784](#)]
3. Johnson RJ, Wolinsky FD. The structure of health status among older adults: disease, disability, functional limitation, and perceived health. *J Health Soc Behav* 1993 Jun;34(2):105-121. [Medline: [8277124](#)]
4. Malmberg JJ, Miilunpalo SI, Vuori IM, Pasanen ME, Oja P, Haapanen-Niemi NA. A health-related fitness and functional performance test battery for middle-aged and older adults: feasibility and health-related content validity. *Arch Phys Med Rehabil* 2002 May;83(5):666-677. [Medline: [11994806](#)]
5. Malmberg J, Miilunpalo S, Pasanen M, Vuori I, Oja P. Characteristics of leisure time physical activity associated with risk of decline in perceived health--a 10-year follow-up of middle-aged and elderly men and women. *Prev Med* 2005 Jul;41(1):141-150. [Medline: [15917005](#)]
6. American College of Sports Medicine, Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, Minson CT, Nigg CR, et al. American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Med Sci Sports Exerc* 2009 Jul;41(7):1510-1530. [Medline: [19516148](#)]
7. Stuart M, Chard S, Benvenuti F, Steinwachs S. Community exercise: a vital component to healthy aging. *Healthc Pap* 2009;10(1):23-28. [Medline: [20057213](#)]

8. US Department of Health and Human Services. 2002. Physical Activity Fundamental To Preventing Disease URL: <http://aspe.hhs.gov/health/reports/physicalactivity/> [accessed 2013-07-20] [WebCite Cache ID 6IG06zoIY]
9. Landi F, Abbatecola AM, Provinciali M, Corsonello A, Bustacchini S, Manigrasso L, et al. Moving against frailty: does physical activity matter? *Biogerontology* 2010 Oct;11(5):537-545. [doi: [10.1007/s10522-010-9296-1](https://doi.org/10.1007/s10522-010-9296-1)] [Medline: [20697813](https://pubmed.ncbi.nlm.nih.gov/20697813/)]
10. Thibaud M, Bloch F, Tournoux-Facon C, Brèque C, Rigaud AS, Dugué B, et al. Impact of physical activity and sedentary behaviour on fall risks in older people: a systematic review and meta-analysis of observational studies. *Eur Rev Aging Phys Act* 2011;8(1):5-15.
11. Hilbe J, Schulz E, Linder B, Them C. Development and alarm threshold evaluation of a side rail integrated sensor technology for the prevention of falls. *Int J Med Inform* 2010 Mar;79(3):173-180. [doi: [10.1016/j.ijmedinf.2009.12.004](https://doi.org/10.1016/j.ijmedinf.2009.12.004)] [Medline: [20083427](https://pubmed.ncbi.nlm.nih.gov/20083427/)]
12. Mei YY, Marquard J, Jacelon C, Defeo AL. Designing and evaluating an electronic patient falls reporting system: Perspectives for the implementation of health information technology in long-term residential care facilities. *Int J Med Inform* 2011 Apr 7:[Epub ahead of print]. [doi: [10.1016/j.ijmedinf.2011.03.008](https://doi.org/10.1016/j.ijmedinf.2011.03.008)] [Medline: [21482183](https://pubmed.ncbi.nlm.nih.gov/21482183/)]
13. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. *JAMA* 2011 Jan 5;305(1):50-58 [FREE Full text] [doi: [10.1001/jama.2010.1923](https://doi.org/10.1001/jama.2010.1923)] [Medline: [21205966](https://pubmed.ncbi.nlm.nih.gov/21205966/)]
14. Ferrucci L, Baldasseroni S, Bandinelli S, de Alfieri W, Cartei A, Calvani D, et al. Disease severity and health-related quality of life across different chronic conditions. *J Am Geriatr Soc* 2000 Nov;48(11):1490-1495. [Medline: [11083330](https://pubmed.ncbi.nlm.nih.gov/11083330/)]
15. Obi T, Ishmatova D, Iwasaki N. Promoting ICT innovations for the ageing population in Japan. *Int J Med Inform* 2013 Apr;82(4):e47-e62. [doi: [10.1016/j.ijmedinf.2012.05.004](https://doi.org/10.1016/j.ijmedinf.2012.05.004)] [Medline: [22727612](https://pubmed.ncbi.nlm.nih.gov/22727612/)]
16. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. *BMC Med Res Methodol* 2010;1:1-10 [FREE Full text] [doi: [10.1186/1471-2288-10-1](https://doi.org/10.1186/1471-2288-10-1)] [Medline: [20053272](https://pubmed.ncbi.nlm.nih.gov/20053272/)]
17. University of Trento. Apple App Store. 2012. IT4Life ActiveLifestyle URL: <https://itunes.apple.com/us/app/it4life-activelifestyle/id465295073?mt=8> [accessed 2013-07-18] [WebCite Cache ID 6IDK7DSTG]
18. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 Feb 27;287(8):1022-1028. [Medline: [11866651](https://pubmed.ncbi.nlm.nih.gov/11866651/)]
19. Eckert JK, Morgan LA, Swamy N. Preferences for receipt of care among community-dwelling adults. *J Aging Soc Policy* 2004;16(2):49-65. [doi: [10.1300/J031v16n02_04](https://doi.org/10.1300/J031v16n02_04)] [Medline: [15148044](https://pubmed.ncbi.nlm.nih.gov/15148044/)]
20. Silveira P, van het Reve E, Daniel F, Casati F, de Bruin ED. Motivating and assisting physical exercise in independently living older adults: a pilot study. *Int J Med Inform* 2013 May;82(5):325-334. [Medline: [23273420](https://pubmed.ncbi.nlm.nih.gov/23273420/)]
21. Mann WC, Ottenbacher KJ, Fraas L, Tomita M, Granger CV. Effectiveness of assistive technology and environmental interventions in maintaining independence and reducing home care costs for the frail elderly. A randomized controlled trial. *Arch Fam Med* 1999;8(3):210-217. [Medline: [10333815](https://pubmed.ncbi.nlm.nih.gov/10333815/)]
22. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. *BMC Med* 2012;10:144 [FREE Full text] [doi: [10.1186/1741-7015-10-144](https://doi.org/10.1186/1741-7015-10-144)] [Medline: [23171675](https://pubmed.ncbi.nlm.nih.gov/23171675/)]
23. Gay V, Leijdekkers P. Personalised mobile health and fitness apps: lessons learned from myFitnessCompanion®. *Stud Health Technol Inform* 2012;177:248-253. [Medline: [22942063](https://pubmed.ncbi.nlm.nih.gov/22942063/)]
24. 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/)]
25. Zwerver J, Weir A, Tol JL. What's app? sports medicine physicians should not talk double Dutch. *Br J Sports Med* 2012 Sep;46(12):833-834. [Medline: [23130364](https://pubmed.ncbi.nlm.nih.gov/23130364/)]
26. 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/)]
27. Consolvo S, Klasnja P, McDonald DW, Landay JA. Goal-setting considerations for persuasive technologies that encourage physical activity. 2009 Presented at: 4th international Conference on Persuasive Technology; Apr 26-29, 2009; Claremont, CA. [doi: [10.1145/1541948.1541960](https://doi.org/10.1145/1541948.1541960)]
28. Lin JJ, Mamykina L, Lindtner S. Fish'n'steps: Encouraging physical activity with an interactive computer game. *Ubiquitous Computing* 2006:261-278.
29. Albaina IM, Visser T, van der Mast C. Flowie: A persuasive virtual coach to motivate elderly individuals to walk. 2009 Presented at: 3rd International Conference on Pervasive Computing Technologies for Healthcare; Apr 1-3, 2009; London p. 1-7.
30. Silveira P, Daniel F, Casati F, van het Reve E, de Bruin ED. International Reports on Socio-Informatics: Designing for Inter/Generational Communities. 2012. ActiveLifestyle: an application to help elderly stay physically and socially active URL: <http://www.iisi.de/fileadmin/IISI/upload/IRSI/IRSI-Vol9-Iss1-2012-finale.pdf> [accessed 2013-07-20] [WebCite Cache ID 6IG1fML2L]
31. LifeParticipation Project. 2012. URL: <http://www.lifeparticipation.org/active.html> [accessed 2013-07-18] [WebCite Cache ID 6IDK45mvN]

32. YouTube. 2012. ActiveLifestyle app video URL: http://www.youtube.com/watch?v=XSFjHcUtt_A&feature=youtu.be [accessed 2013-07-18] [WebCite Cache ID 6IDK5sKzc]
33. Cress ME, Buchner DM, Prohaska T, Rimmer J, Brown M, Macera C, et al. Best practices for physical activity programs and behavior counseling in older adult populations. *J Aging Phys Act* 2005 Jan;13(1):61-74. [Medline: [15677836](#)]
34. Kraemer WJ, Ratamess NA. Fundamentals of resistance training: progression and exercise prescription. *Med Sci Sports Exerc* 2004 Apr;36(4):674-688. [Medline: [15064596](#)]
35. YouTube. 2012. LifeParticipation Project: ActiveLifestyle videos URL: http://www.youtube.com/user/psilvieraf81?feature=results_main [accessed 2013-07-18] [WebCite Cache ID 6IDKFDTwV]
36. YouTube. 2012. Longer study of the Active Lifestyle app in Switzerland URL: <http://www.youtube.com/watch?v=mNaIOYqN6tU&feature=youtu.be> [accessed 2013-07-18] [WebCite Cache ID 6IDKGyCew]
37. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000 Sep 16;321(7262):694-696 [FREE Full text] [Medline: [10987780](#)]
38. Arikawa AY, O'Dougherty M, Schmitz KH. Adherence to a strength training intervention in adult women. *J Phys Act Health* 2011 Jan;8(1):111-118 [FREE Full text] [Medline: [21297191](#)]
39. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. *Age Ageing* 2012 Jan;41(1):16-23 [FREE Full text] [doi: [10.1093/ageing/afr103](#)] [Medline: [21875865](#)]
40. de Bruin ED, Schmidt A. Walking behaviour of healthy elderly: attention should be paid. *Behav Brain Funct* 2010;6:59 [FREE Full text] [doi: [10.1186/1744-9081-6-59](#)] [Medline: [20939911](#)]
41. McDonough AL, Batavia M, Chen FC, Kwon S, Ziai J. The validity and reliability of the GAITRite system's measurements: A preliminary evaluation. *Arch Phys Med Rehabil* 2001 Mar;82(3):419-425. [doi: [10.1053/apmr.2001.19778](#)] [Medline: [11245768](#)]
42. Hartmann A, Luzi S, Murer K, de Bie RA, de Bruin ED. Concurrent validity of a trunk tri-axial accelerometer system for gait analysis in older adults. *Gait Posture* 2009 Apr;29(3):444-448. [doi: [10.1016/j.gaitpost.2008.11.003](#)] [Medline: [19070494](#)]
43. Prochaska JO, Butterworth S, Redding CA, Burden V, Perrin N, Leo M, et al. Initial efficacy of MI, TTM tailoring and HRI's with multiple behaviors for employee health promotion. *Prev Med* 2008 Mar;46(3):226-231 [FREE Full text] [doi: [10.1016/j.ypmed.2007.11.007](#)] [Medline: [18155287](#)]
44. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: L Erlbaum Associates; 1988.
45. Gibson MJ, Andres RO, Isaacs B. The prevention of falls in later life. A report of the Kellogg International Work Group on the Prevention of Falls by the Elderly. *Dan Med Bull* 1987 Apr;34 Suppl 4:1-24. [Medline: [3595217](#)]
46. Yardley L, Donovan-Hall M, Francis K, Todd C. Attitudes and beliefs that predict older people's intention to undertake strength and balance training. *J Gerontol B Psychol Sci Soc Sci* 2007 Mar;62(2):P119-P125. [Medline: [17379672](#)]
47. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. *Patient Educ Couns* 2013 Apr;91(1):14-24. [doi: [10.1016/j.pec.2012.10.018](#)] [Medline: [23194823](#)]
48. Moschny A, Platen P, Klaassen-Mielke R, Trampisch U, Hinrichs T. Barriers to physical activity in older adults in Germany: a cross-sectional study. *Int J Behav Nutr Phys Act* 2011;8:121 [FREE Full text] [doi: [10.1186/1479-5868-8-121](#)] [Medline: [22047024](#)]
49. Ostir GV, Berges I, Kuo YF, Goodwin JS, Ottenbacher KJ, Guralnik JM. Assessing gait speed in acutely ill older patients admitted to an acute care for elders hospital unit. *Arch Intern Med* 2012 Feb 27;172(4):353-358 [FREE Full text] [doi: [10.1001/archinternmed.2011.1615](#)] [Medline: [22371922](#)]
50. Studenski S. Bradypedia: is gait speed ready for clinical use? *J Nutr Health Aging* 2009 Dec;13(10):878-880. [Medline: [19924347](#)]
51. Hardy SE, Perera S, Roumani YF, Chandler JM, Studenski SA. Improvement in usual gait speed predicts better survival in older adults. *J Am Geriatr Soc* 2007 Nov;55(11):1727-1734. [doi: [10.1111/j.1532-5415.2007.01413.x](#)] [Medline: [17916121](#)]
52. Anderson-Bill ES, Winett RA, Wojcik JR. Social cognitive determinants of nutrition and physical activity among web-health users enrolling in an online intervention: the influence of social support, self-efficacy, outcome expectations, and self-regulation. *J Med Internet Res* 2011;13(1):e28 [FREE Full text] [doi: [10.2196/jmir.1551](#)] [Medline: [21441100](#)]
53. Bhardwaj SS, Camacho F, Derrow A, Fleischer AB, Feldman SR. Statistical significance and clinical relevance: the importance of power in clinical trials in dermatology. *Arch Dermatol* 2004 Dec;140(12):1520-1523. [doi: [10.1001/archderm.140.12.1520](#)] [Medline: [15611433](#)]

Abbreviations

ANOVA: analysis of variance

ETH: Eidgenössische Technische Hochschule

IT: information technology

TTM: Transtheoretical Model

Edited by G Eysenbach; submitted 14.02.13; peer-reviewed by MC Costabile, E Stanmore; comments to author 21.03.13; revised version received 25.04.13; accepted 08.06.13; published 12.08.13.

Please cite as:

Silveira P, van de Langenberg R, van het Reve E, Daniel F, Casati F, de Bruin ED

Tablet-Based Strength-Balance Training to Motivate and Improve Adherence to Exercise in Independently Living Older People: A Phase II Preclinical Exploratory Trial

J Med Internet Res 2013;15(8):e159

URL: <http://www.jmir.org/2013/8/e159/>

doi: [10.2196/jmir.2579](https://doi.org/10.2196/jmir.2579)

PMID: [23939401](https://pubmed.ncbi.nlm.nih.gov/23939401/)

©Patrícia Silveira, Rolf van de Langenberg, Eva van het Reve, Florian Daniel, Fabio Casati, Eling D de Bruin. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 12.08.2013. 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

Standardized, App-Based Disinfection of iPads in a Clinical and Nonclinical Setting: Comparative Analysis

Urs-Vito Albrecht^{1*}, Dr med, MPH; Ute von Jan^{1*}, Dr rer biol hum; Ludwig Sedlacek^{2*}, Dr med; Stephanie Groos^{3*}, Dr med; Sebastian Suerbaum^{2*}, Dr med; Ralf-Peter Vonberg^{2*}, Dr med

¹PL Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

²Institute for Medical Microbiology and Hospital Epidemiology, Hannover Medical School, Hannover, Germany

³Institute of Cell Biology in the Center of Anatomy, Hannover Medical School, Hannover, Germany

* all authors contributed equally

Corresponding Author:

Urs-Vito Albrecht, Dr med, MPH
PL Reichertz Institute for Medical Informatics
Hannover Medical School
Carl-Neuberg-Str 1
Hannover, 30625
Germany
Phone: 49 511 532 ext 3508
Fax: 49 511 532 2517
Email: albrecht.urs-vito@mh-hannover.de

Abstract

Background: With the use of highly mobile tools like tablet PCs in clinical settings, an effective disinfection method is a necessity. Since manufacturers do not allow cleaning methods that make use of anything but a dry fleece, other approaches have to be established to ensure patient safety and to minimize risks posed by microbiological contamination.

Objective: The ability of isopropanol wipes to decontaminate iPads was evaluated prospectively in an observer blinded, comparative analysis of devices used in a clinical and a nonclinical setting.

Methods: 10 new iPads were randomly deployed to members of the nursing staff of 10 clinical wards, to be used in a clinical setting over a period of 4 weeks. A pre-installed interactive disinfection application (deBac-app, PLRI MedAppLab, Germany) was used on a daily basis. Thereafter, the number and species of remaining microorganisms on the surface of the devices (13 locations; front and back) was evaluated using contact agar plates. Following this, the 10 iPads were disinfected and randomly deployed to medical informatics professionals who also used the devices for 4 weeks but were forbidden to use disinfecting agents. The quality of a single, standardized disinfection process was then determined by a final surface disinfection process of all devices in the infection control laboratory. No personal data were logged with the devices. The evaluation was performed observer blinded with respect to the clinical setting they were deployed in and personnel that used the devices.

Results: We discovered a 2.7-fold (Mann-Whitney U test, $z=-3.402$, $P=.000670$) lower bacterial load on the devices used in the clinical environment that underwent a standardized daily disinfection routine with isopropanol wipes following the instructions provided by “deBac-app”. Under controlled conditions, an average reduction of the mainly Gram-positive normal skin microbiological load of 99.4% (Mann-Whitney U test, $z=-3.1798$, $P=.001474$) for the nonclinical group and 98.1% (Mann-Whitney U test, $z=3.1808$, $P=.001469$) for the clinical group was achieved using one complete disinfecting cycle.

Conclusions: Normal use of tablet PCs leads to a remarkable amount of microbial surface contamination. Standardized surface disinfection with isopropanol wipes as guided by the application significantly reduces this microbial load. When performed regularly, the disinfection process helps with maintaining a low germ count during use. This should reduce the risk of subsequent nosocomial pathogen transmission. Unfortunately, applying a disinfection procedure such as the one we propose may lead to losing the manufacturer’s warranty for the devices; this remains an unsolved issue.

(*J Med Internet Res* 2013;15(8):e176) doi:[10.2196/jmir.2643](https://doi.org/10.2196/jmir.2643)

KEYWORDS

tablet PC; hygiene; disinfection; hygiene; nosocomial transmission

Introduction

Infections are called nosocomially acquired if they occur during a hospital stay. They have an enormous clinical and economical impact for health care systems [1]. In addition, multidrug resistant pathogens represent an increasing problem in hospitals these days [2]. Besides the hands of health care workers (HCW), contaminated medical devices and surfaces play an important role in the transmission of bacterial pathogens. This necessitates considerable effort for environmental infection control in order to prevent the spread of all kinds of microorganisms between patients [3].

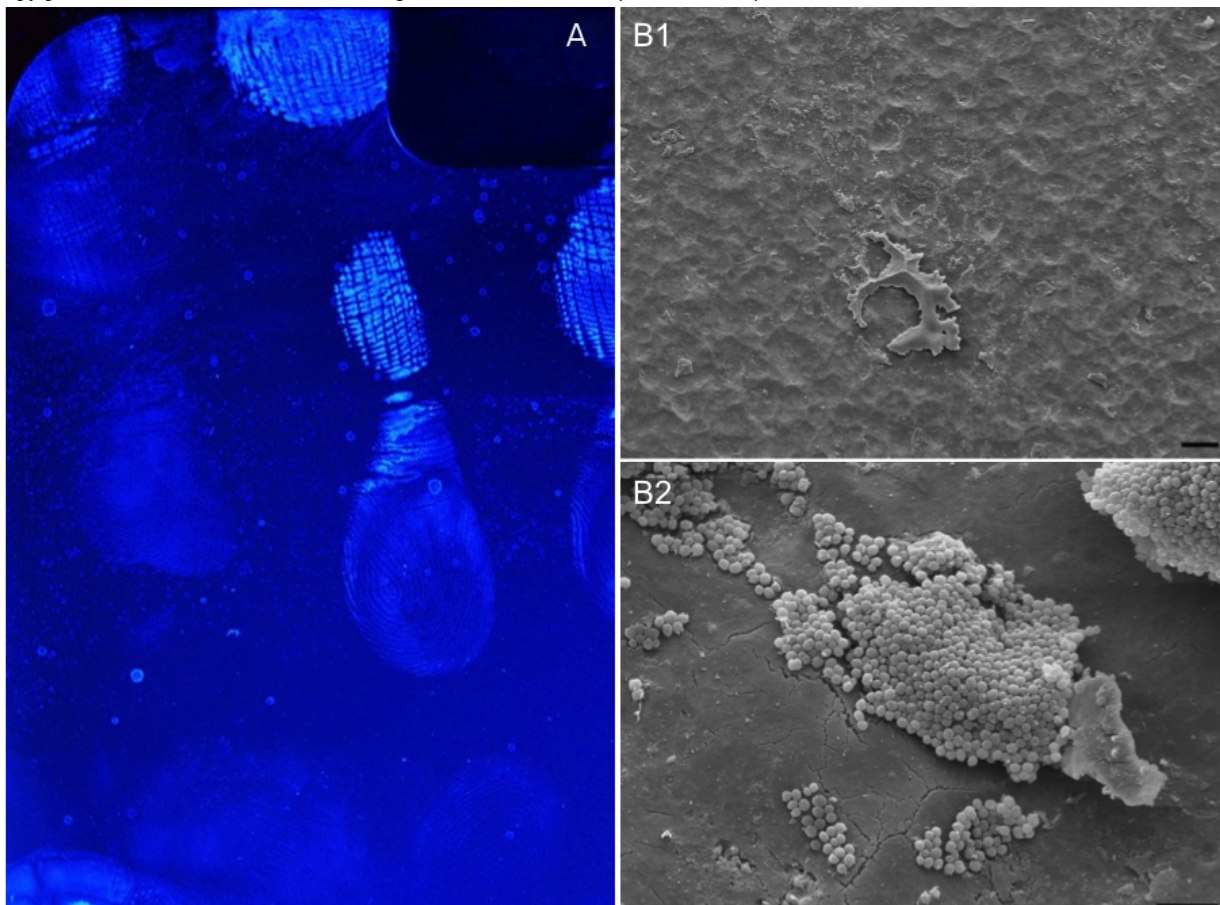
Mobile devices such as mobile phones or personal digital assistants (PDAs) represent a rather novel “surface” in the hospital setting that may also play an important role in the transmission of nosocomial pathogens. Nowadays, such devices are frequently used by physicians and other medical staff for both clinical practice and educational purposes [4,5]. The number and availability of medical applications (apps) on smartphones is constantly rising and includes drug guides, medical calculators, coding and billing apps, textbooks and other reference materials, classification and treatment algorithms, as well as information regarding general medical knowledge [6]. However, contamination of a device’s surface

occurs every time it is being touched by a user [7], and there are several reports showing that these devices may then serve as vectors for transmission of pathogens to patients [8,9]. In a review of data from studies published between 2002 and 2008, Brady et al [10] showed that 9-25% of mobile communication devices were contaminated with pathogenic bacteria. More recent prevalence studies report contamination rates as high as 44-95% [7,11-15]. Like mobile phones, tablet PCs (for example the iPad) are also frequently touched during patient care and bacteria may reside on their surfaces (Figure 1).

Before our study began, tests conducted with microbiological swabs showed that brand new iPad devices are not significantly contaminated with bacteria or fungi. However, the extent of contamination of tablets that have already been used remains yet unknown. Considering that bacteria may survive for days and weeks on inanimate surfaces [16], there is a need to determine the extent of contamination and to implement proper routine decontamination measures.

The present study was set up to determine (1) the microbiological flora (qualitative and quantitative) on tablet PCs as a result of use under the usual conditions that can be found in clinical as well as in nonclinical settings and (2) the quality of a standardized disinfection process as guided by an app specifically programmed for this purpose.

Figure 1. Aluminium backside of a tablet PC with fingerprints and other residue visible under fluorescent light and corresponding scanning electron microscopy pictures of cocci on the device in 2 magnifications (Bars: B1=5µm and B2=10µm).



Methods

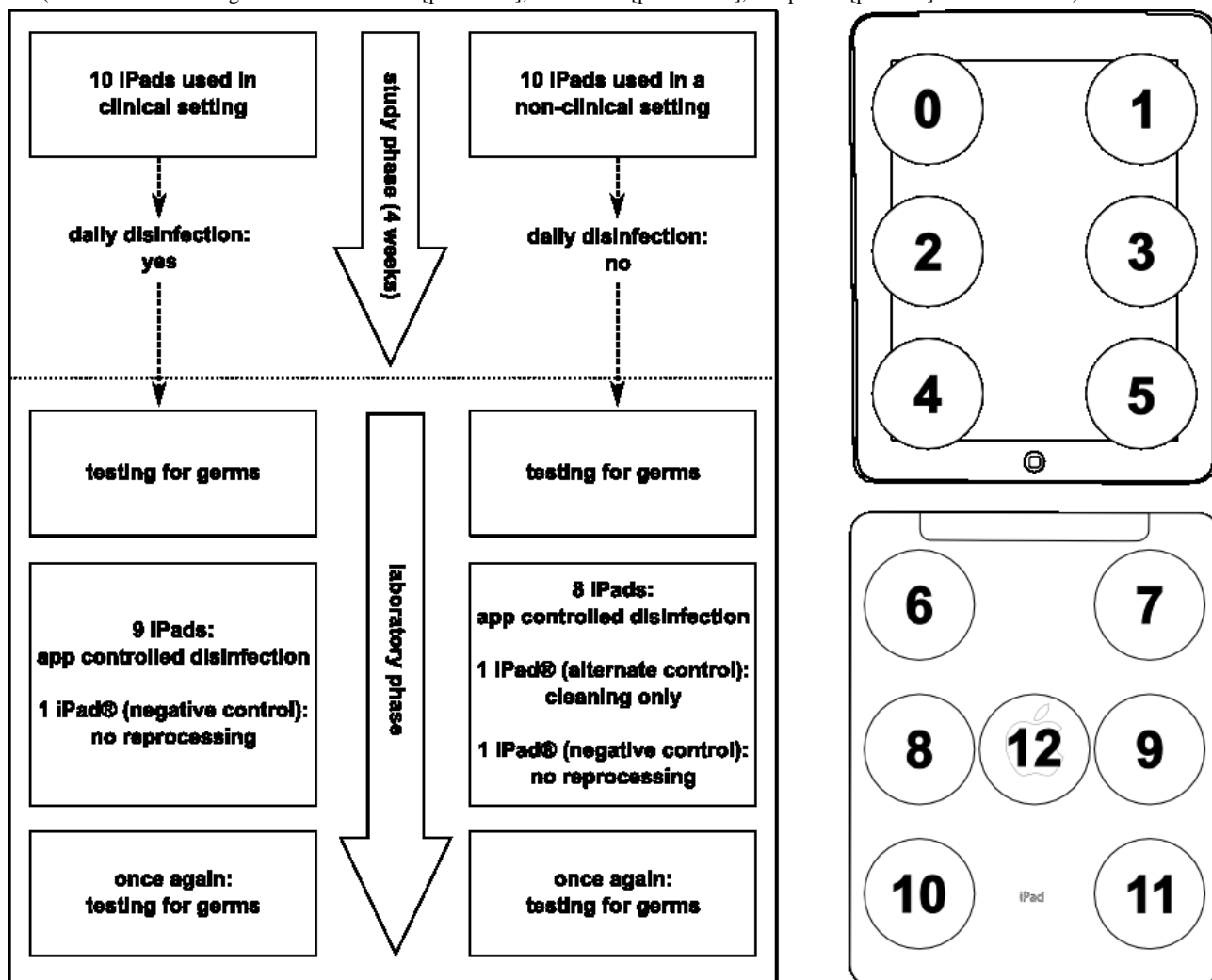
Clinical Setting

A set of 10 Apple iPads was randomly distributed to be used by nursing staff after obtaining informed consent. Various wards of Hannover Medical School, a tertiary care German university hospital, were included. They covered nonsurgical as well as surgical specialties. For the clinical setting, the disinfection study was an add-on to a larger trial dealing with various aspects [17] of using iPads on the wards of the Hannover Medical School. Altogether, approximately 160 staff members on the wards had the opportunity to use the devices. It was not possible to determine how many different individuals had used the iPads. As mentioned in [17], the return rate for questionnaires dealing with the overall project was approximately 26% (42/160); this can be assumed as the minimum of actual individual users. Regarding age and gender, the demographics of the participants who had returned the questionnaire paralleled the values for the nursing staff at the Hannover Medical School, where, at the end of 2012, 83% out of the total 2596 employees of the nursing staff were female (data obtained from the human resources department). For our study, 85% (36/42) were female and the age distribution of 69% (29/42) for those below 45 years of age

and 31% (13/42) for those 45 years of age or older was also comparable. It was not possible to determine how they had used the devices and whether all or only some of them had disinfected the devices aided by deBac-app.

There were no additional accessories such as protective cases, polyurethane foils, rubber or silicone covers, since these may add additional, hard-to-disinfect niches for contaminating pathogens. Instead, we recommended disinfection of the plain surface of every device on a daily basis: once at the beginning of the working shift as well as anytime when an obvious contamination had occurred. Standardization of the disinfection process was achieved by using the “deBac-app” (from the MedAppLab, Hannover, Germany), which was preinstalled on all iPads. This app was designed by our research group and is an interactive cleaning and disinfection guide that is available free from Apple’s App Store [18]. It provides users with simple-to-follow instructions on how to properly disinfect the entire device (Multimedia Appendix 1). Every disinfection process was logged locally on the respective device. No personal data were stored or transmitted to the observers. The study time period of this clinical setting study arm was set to 4 weeks between August 28 and September 19, 2011, in order to achieve a steady state in terms of usage and reprocessing. A flowchart of the study’s timeline is shown in Figure 2.

Figure 2. Flowchart on the timeline of the 2 settings of iPad usage (clinical and nonclinical) and contact points used for microbiological sampling of an iPad (surface material was glass on the front side [points 1-6], aluminium [points 7-12], and plastic [point 13] on the backside).



Nonclinical Setting

10 iPads were also used for a 4-week time period, between September 23 and October 21, 2011, by 10 members of the medical information technology staff (30% or 3/10 female, aged 23–63 years, mean 41.7, SD 10.6) in the nonclinical study arm. The devices were randomly distributed to the staff after obtaining informed consent for participation in the study. As these staff members generally do not have contact to patients, no regular disinfection of the tablet PCs was performed (Figure 2). All 10 participants of the nonclinical part of the study belonged to the P.L. Reichertz Institute for Medical Informatics at the Hannover Medical School and had volunteered to participate. For the nonclinical setting, there were no dropouts during the course of the study and all participants stated that they had frequently used their devices during the 4-week period.

Microbiological Testing

After 4 weeks of usage, all devices were examined for growth of microorganisms on their surfaces as soon as they arrived at the microbiological laboratory. Culture media with a contact area of 25 cm² (CASO contact agar plates, Heipha diagnostica Dr. Müller GmbH) were used. These culture media support growth of most aerobic bacteria, molds, and yeasts. Since it is known that the adherence and survival of microorganisms may vary depending upon the type of surfaces material [15], we decided to perform the microbiological sampling for 13 different contact points of each iPad (Figure 2, right), including locations on the front and the back side of the device. These contact points covered all types of material to be found on the surface of the devices (glass, plastic, and aluminium). The contact plates were incubated at 37°C for 18 hours under aerobic conditions. Colony forming units (CFU) were then counted, and species differentiation was carried out in the microbiology laboratory of our facility according to the national guidelines of the German Institute for Standardization DIN EN ISO 15189 as certified by the German Accreditation Council (DAR). The evaluation was conducted in such a way that the laboratory was unaware of the setting to which the iPads had been deployed.

Electron Microscopy

For photo documentation of bacterial contamination on the tablet PC's surfaces, scanning electron microscopy (SEM) was applied (Figure 1). Specimens were fixed in 0.1M Na-Cacodylate-HCl buffer (pH 7.3) containing 3% glutaraldehyde for at least 4 hours at 4°C. After washing in the buffer of the fixative, the cells were postfixed in 2% OsO₄ buffered in Na-Cacodylate 0.1 M for 90 minutes at room temperature, dehydrated in ascending concentrations of acetone, and subsequently dried in a Balzers CPD 030 critical point dryer (Bal-Tec-AG). After mounting on aluminium stubs with conductive carbon cement (Plano) and sputter coating with gold in a Polaron E 5400 sputter coater, the samples were investigated in a Philips SEM 505 scanning electron microscope at an acceleration voltage of 10kV. Images were recorded using the SEM software version 2.0 [19].

Final Reprocessing

After primary microbiological testing, the devices underwent final reprocessing performed by laboratory staff (Figure 2). For

the clinical setting, with one exception, all devices were disinfected using isopropanol wipes (mikrozid-AF, Schülke & Mayr GmbH) using the 6-step disinfection process guided by deBac-app as it was described above (Multimedia Appendix 1). The remaining 10th iPad did not get reprocessed and, thus, served as a negative control for the disinfection process. For the 10 iPads that had been used in the nonclinical setting, we chose a slightly different approach: 8 of them were disinfected as described, while 1 remained without treatment and 1 was simply cleaned (but not disinfected) by using a new “soft, lint-free cloth”, without any liquid cleaning agents, as recommended in the instructions of the manufacturer of the iPad [20]. A final microbiological testing as described above was performed following the different, aforementioned types of reprocessing performed in the laboratory (Figure 2).

Nasal Swabs

People may be physiologically colonized by *Staphylococcus aureus* in the anterior nose, and some of these strains even show multidrug resistance, so-called methicillin-resistant *Staphylococcus aureus* (MRSA) [21]. For a comparison of the *Staphylococcus aureus* colonization status of the 10 medical informatics professionals and the surface of their devices, nasal swabs (Transystem, Lot 9275, Hain Lifescience) were taken from the users after informed consent was obtained. Swabs were cultured on Columbia 5% sheep blood agar (Becton, Dickinson) overnight at 37°C. Species identification and susceptibility testing were then performed according to laboratory standard operation protocols.

Statistical Analysis

We expected a very strong effect of the applied method on reduction of the CFU according to the literature [22]. Therefore, a smaller sample size was expected to be sufficient to demonstrate the efficacy of the disinfection. Calculated for a paired nonparametric test [23], a sample size of 6 iPads per group was considered sufficient to show a significant effect of a reduction of 98% with $\beta=0.20$. The sample size calculation was performed with nQuery Advisor V.7, Statistical Solutions. Since a normal distribution of bacteria on the devices and sampled locations could not be confirmed by descriptive statistics, the Wilcoxon signed-rank test and the Mann-Whitney U test were applied (IBM SPSS Statistics Version 20). All tests were performed two-sided, with $\alpha=0.05$. Microsoft Excel 2007 was used for qualitative descriptive and quantitative data analysis. Intraclass correlation could not be confirmed following Shrout and Fleiss' two-way random single measures (consistency) approach computed with SPSS [24].

Results

Qualitative and Quantitative Analysis of Microbiological Flora

A total of 6811 colonies representing microbial growth were detected during the initial testing of the iPads after use: 1842 CFU on tablet PCs from the hospital wards where the devices had been disinfected regularly, compared to 4,969 CFU recovered from tablet PCs from the nonclinical setting where daily disinfections had not been carried out (Mann-Whitney U

test, $z=-3.402$, $P=.000670$). The distribution of pathogens on the various parts of the tablet PCs is shown in Table 1. Note that most pathogens found in both study arms (clinical and nonclinical) were gram-positive bacteria. A more detailed analysis of the species is shown in Table 2. The majority of microorganisms were members of the physiological microbiota of the human skin. The distribution did not differ significantly between both study arms. The main bacterial genera grown from iPads from the clinical setting were staphylococci ($n=1104$; 59.9%) and micrococci ($n=469$; 25.7%). The same types of bacteria were mainly found on iPads from the nonclinical setting (staphylococci: $n=3678$; 74.2% and micrococci: $n=1051$; 21.2%). However, the cultured microorganisms also included several pathogens. The most frequently identified pathogen was *Staphylococcus aureus* (non-MRSA only), which was found in nasal swabs from 2 medical informatics professionals as well as on their tablet PCs, but this species was also detected on tablet PCs from 2 other staff members who were not colonized themselves (Wilcoxon signed-rank test, $z=-1.414$, $P=.157$).

Determination of the Quality of the Guided Standardized Disinfection Process

The percentage of reduction of pathogens on iPads that underwent the standardized disinfection protocol is shown in Table 3 and Figures 3 and 4. A significant overall reduction on microbes was achieved for both the clinical setting (98.1%; Wilcoxon signed-rank test, $z=-3.1808$, $P=.001469$) and the nonclinical setting (99.4%; Mann-Whitney U test, $z=-3.1798$; $P=.001474$). Note that bacilli are capable of forming spores. If

doing so, those spores then show an extraordinarily high resistance towards disinfection processes (as they cannot at all be inactivated by alcohol-based disinfection) and other extreme environmental effects [25]. Still, a reduction of *Bacillus* spp. of 88% was achieved in our study in both settings. The reduction rates of all other bacterial and fungal species were as high as 99%.

Re-Sampling of the Two Untreated (Control) iPads

As mentioned above and shown in Figure 2, two tablet PCs were sampled once again without any additional reprocessing step in order to check for a potential germ-reducing effect due to the first sampling process itself. Re-sampling revealed reduction rates of 11.4% (228 of 246 CFU) on the nondisinfected iPad from the clinical setting and 22.4% (595 of 767 CFU) on the nondisinfected iPad from nonclinical setting only.

Determination of the Quality of Cleaning With a Soft, Lint-Free Cloth Without Liquid Cleaning Agents

As already noted (see Figure 2), one tablet PC from the nonclinical study was cleaned only with a brand-new fleece according to the instructions of the manufacturer. The initial CFU count of this device at arrival in the laboratory was 891 CFU; 427 CFU remained after cleaning with a fleece (reduction rate: 51.1%). Removal of bacteria was rather higher for the glass surface of the front (231 of 234 CFU; 98.7% reduction), but almost no reduction (5 of 77 CFU; reduction rate: 6.5%) was achieved on the plastic part of the device. Cleaning the aluminium resulted in a CFU reduction from 580 to 352 corresponding to 38.3%.

Table 1. Recovery of pathogens found on the devices' surfaces on initial arrival at the laboratory (shown as cumulative number of CFU from 10 tablet PCs each). Comparison of the total number of microorganisms: Mann-Whitney U test, $z=-3.402$; $P=.000670$.

	Clinical setting			Nonclinical setting		
	Total CFU	Median CFU	IQR	Total CFU	Median CFU	IQR
Total	1842	162	125.75	4969	440	273.75
Gram positive bacteria	1825	160.5	122.75	4916	437.5	283
Front (glass)	772	58.5	62.25	1,672	167	104.25
Back (plastic)	214	22.5	27	481	46	35.5
Back (aluminium)	839	63	68	2763	300.5	183.25
Gram negative bacteria	9	1	0.75	52	2	4.5
Front (glass)	6	1	1	17	0.5	3.75
Back (plastic)	0	0	0	5	0	1
Back (aluminium)	3	0	0.75	30	0	1
Other	8	0	1.5	1	0	0
Front (glass)	6	0	1.5	1	0	0
Back (plastic)	1	0	0	0	0	0
Back (aluminium)	1	0	0	0	0	0

Table 2. Distribution of species of microorganisms from the surface of the iPads on initial arrival of the devices at the laboratory (shown as cumulative number of colony forming units from 20 tablet PCs; n=6811).

	CFU	%	Gram stain
Physiological human skin flora			
<i>Staphylococcus epidermidis</i>	1783	26.2	positive
<i>Micrococcus luteus</i>	1509	22.2	positive
<i>Staphylococcus hominis</i>	1256	18.4	positive
<i>Staphylococcus capitis</i>	977	14.3	positive
<i>Staphylococcus warneri</i>	194	2.9	positive
Other coagulase-negative staphylococci	363	5.3	positive
<i>Bacillus</i> spp.	309	4.5	positive
<i>Corynebacterium</i> spp.	117	1.7	positive
Other species	20	0.3	positive
Pathogenic microorganisms			
<i>Staphylococcus aureus</i> (non-MRSA ^a)	218	3.2	positive
<i>Pseudomonas</i> spp.	36	0.5	negative
<i>Aspergillus</i> spp. / molds	9	0.1	N/A
<i>Acinetobacter</i> spp.	8	0.1	negative
Other species	12	0.2	negative

^aMRSA: methicillin resistant *Staphylococcus aureus*.

Figure 3. Histogram of CFU-count per localization samples taken from 6 corresponding devices in a clinical and nonclinical setting, stratified for position number, side, and material.

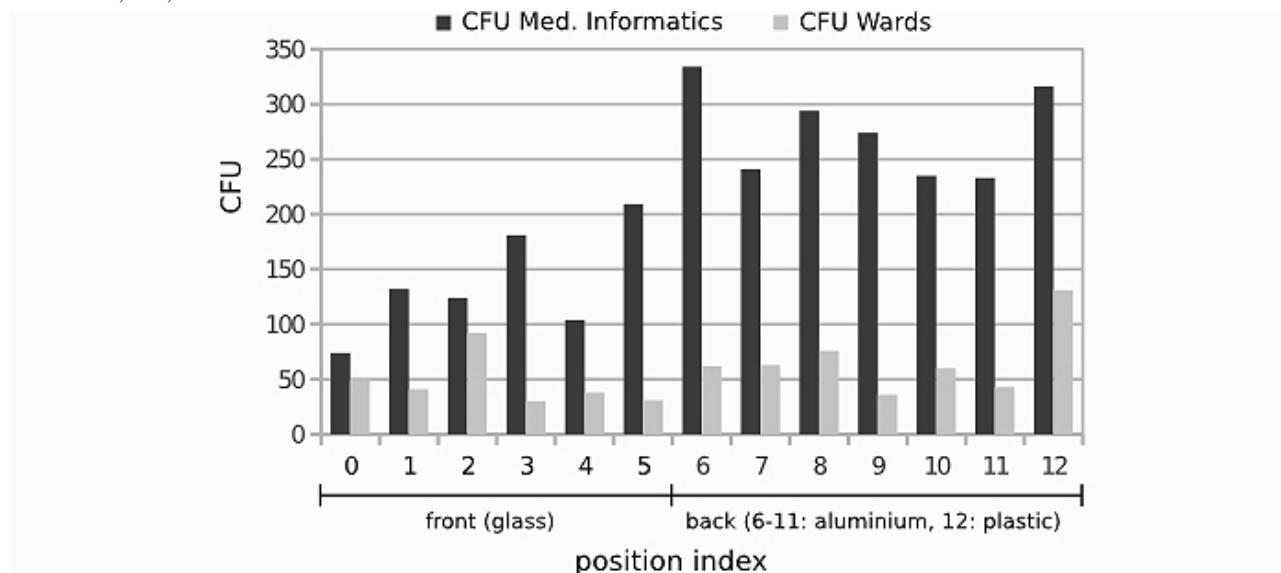


Figure 4. Reduction of CFU in percent per position, side, and material after disinfection.

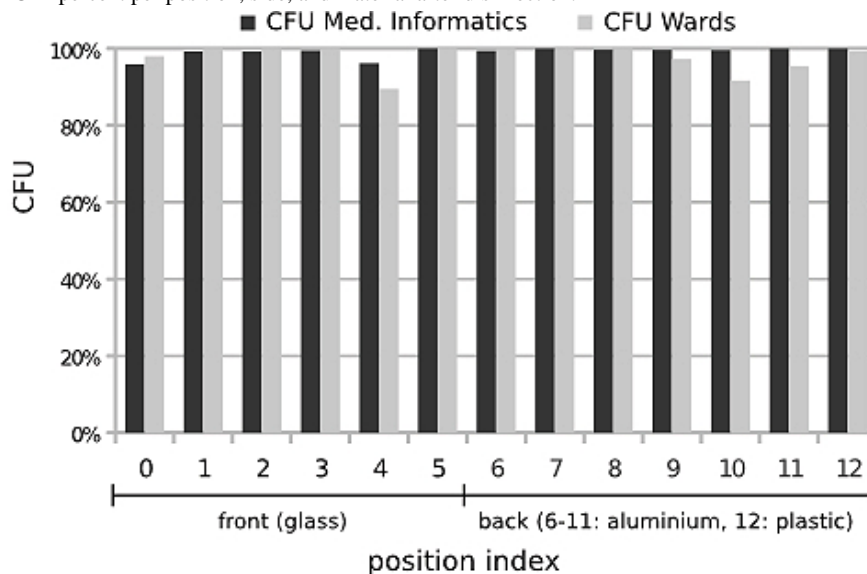


Table 3. Reduction of bacteria on the surface of 6 iPads after standard disinfection procedure stratified by the type of previous usage (clinical vs nonclinical), the sample location (front vs back), the type of material (glass vs aluminium vs plastic), and type of Gram stain (positive vs negative).

	On laboratory arrival			After standardized disinfection			
	Total CFU	Median CFU	IQR	Total CFU	Median CFU	IQR	CFU reduction, %
Total (clinical setting)	753	121	65.75	14	2	1.5	98.1
Gram positive bacteria	749	121	65	14	2	1.5	97.2
Front (glass)	280	39.5	42.75	5	1	0.75	96.9
Back (plastic)	131	18	32.75	1	0	0	98.5
Back (aluminium)	338	51	37.25	8	0.5	1.75	96.5
Gram negative bacteria	4	0.5	1	0	0	0	100.0
Front (glass)	2	0	0.75	0	0	0	100.0
Back (plastic)	0	0	0	0	0	0	100.0
Back (aluminium)	2	0	0.75	0	0	0	100.0
Total (nonclinical setting)	2751	440	81	15	1	4.25	99.4
Gram positive bacteria	2739	437.5	78.25	15	1	5.25	99.5
Front (glass)	816	148.5	107.5	10	1	2.75	97.9
Back (plastic)	315	56.5	32	0	0	0	100.0
Back (aluminium)	1608	300.5	90	5	0	0	99.7
Gram negative bacteria	12	1.5	3.25	0	0	0	100.0
Front (glass)	8	0.5	2.5	0	0	0	100.0
Back (plastic)	1	0	0	0	0	0	100.0
Back (aluminium)	3	0	0.75	0	0	0	100.0

Discussion

Principal Findings

Without any doubt, mobile devices provide numerous advantages in a hospital setting, but despite these benefits, the potential risk of pathogen transmission must be taken into account [10]. There are several conclusions that can be drawn from the set-up and from the findings of our study.

As has been shown for other mobile devices [8,9], an extensive surface contamination also takes place when iPads are being used. Every fingerprint on the surface (Figure 1) will leave residue on the glass, aluminum, and plastic parts of the device (Figure 1) and may contain a large number of bacteria. An increased awareness of this fact is required when those devices are used during patient care. Brady et al [26] questioned 90 HCW (surgeons, anesthesiologists, and medical students)

regarding this issue. At least 53% of them carried one mobile device (16% carried even more than one) including PDAs, mobile phones, and pagers. When asked about their cleaning habits, the HCW admitted that 80% of the PDAs, 85% of the mobile phones, and 96% of the pagers had never been cleaned by the owner.

Most of the pathogens are members of the resident or transient flora of its user (skin and/or anterior nose). Whatever microorganisms are present on the hands will be found on the mobile phone [27,28] or the tablet PC later on. This stresses the need for proper hand hygiene of HCW as it has been addressed by the World Health Organization in the international “clean hands campaigns” recently [29]. Patients, too, should be educated about the role of their own mobile devices brought to the hospital because these devices will also become contaminated [30]. Especially patients who harbor multidrug resistant bacteria should be discouraged to share their mobile phone with others.

As shown by the repeated sampling by contact plates described above, microorganisms may easily spread from the surface of the tablet PC when touched again. A transmission of pathogens that have caused nosocomial outbreaks has been shown for mobile phones [8,12]. One would assume that a much larger device such as a tablet PC is even more likely to serve as a vehicle of infectious agents. HCW should therefore be encouraged to perform alcohol-based hand rubs after using their mobile device [31].

Cleaning with a fleece as recommended by the manufacturer of the tablet PC showed a reduction of about 50% of microorganisms. However, a sufficient reduction of the microbiological load will be achieved only when proper disinfection is performed. A cleaning phase for visible contamination and a disinfection phase as a final decontamination step are considered most effective according to infection control as recommended in the guideline for environmental cleaning in health care facilities from the Centers for Disease Control and Prevention and the Healthcare Infection Control Practices Advisory Committee [3].

As has been shown for the disinfection of mobile phones [32], a disinfection procedure for iPads that makes use of isopropanol wipes is very effective in reducing and inactivating residual bacteria. However, one has to keep in mind that this procedure may cause a loss of the warranty for this product. It is noteworthy that Apple’s recommendations for the cleaning process of the iPad, available on the company’s website, have significantly changed in the past. The version from December 15, 2010, stated that “it is also safe to use isopropyl alcohol 70% or a similar product” for this purpose. However, in the meantime, this statement has been withdrawn. Instead, it is now specified that “liquid damage is not covered under the Apple product warranty or AppleCare Protection Plans” and specifically to “Avoid getting moisture in openings. Don’t use window cleaners, household cleaners, aerosol sprays, solvents, alcohol, ammonia, or abrasives to clean the iPad” [20].

Regular disinfection serves to maintain a significantly lower load of pathogens. Our study results imply that disinfection followed by the deBac-app has the ability to reduce

microbiological flora in a quality manner. We would like to recommend using a standardized scheme for the disinfection process as described and controlled by the deBac application. Reprocessing of tablet PCs should be performed at least once a day, preferably at the beginning of the working shift. Additional courses of disinfection should also be carried out any time that visible contamination has occurred. Furthermore, we recommend disinfecting the device after using it in a patient room under isolation precautions (eg, if the patient harbors some type of multidrug resistant organism). The guided disinfection procedure ensures that all parts of the surface get thoroughly treated. Furthermore, all steps of reprocessing are documented and may be filed in infection control records.

Limitations

Regarding the study design, we were willing to accept the following limitations: in accordance with our in-house regulations, machinery used in a clinical environment has to be disinfected. Therefore, it was impossible to learn about the baseline colonization by installing a control group on the wards and allowing this control group to use the devices without any disinfection. A randomized controlled trial or a controlled design with carefully matched comparison groups using standard practices as compared to the deBac condition would allow verification of the assumption that, when using mobile devices such as iPads in a clinical environment, performing an app-based disinfection process is more effective in reducing microbiological flora than simply using regular hand hygiene. This will have to be addressed in forthcoming studies.

The nursing staff was not provided with additional, paper-bound cleaning instructions since we wanted users to refer to the information available on the devices. We also refrained from collecting any personal data from the devices since our in-house data protection policy had to be followed. Only the information available from the anonymous cleaning protocols acquired from within the app was used for the evaluation and process documentation. The entries found in the log files demonstrated daily usage, but it is unknown to the observers how often and in which way the machines were used when the deBac-app was not running. This, of course, may bias the amount of CFU that were found on the surfaces.

As for the sensitivity and specificity of the various microbiological tests conducted, all laboratory methods that were used during the course of the study, including all microbiological tests for identification of pathogens, have been certified according to national guidelines. However, the sensitivity and specificity of taking the samples remains unclear as there is no so-called “gold standard” to compare with. It is known from optimized protocols that recovery may come up with a sensitivity of 98% and a specificity for a particular pathogen of interest of 95% [33]. However, standard environmental sampling still remains an unsolved problem these days [34,35]. Unfortunately, it is impossible compare the *Staphylococcus aureus* strains cultured from the nasal swabs of the users with those found on the devices themselves since the isolates had been discarded in the meantime. However, to us, it seems highly probable that we found corresponding strains here, as *Staphylococcus aureus* does not represent a typical

“environmental” bacterium. Furthermore, it is well known that people tend to frequently touch their noses.

Conclusions

Cleaning the devices with disinfecting wipes can be considered efficient and effective. Nevertheless, one must be aware of the potential danger of damaging the devices: there will definitely be a breach in warranty if liquid seeps into the device in any way and causes damage. On the other hand, although tablet PCs were originally developed for the consumer market, once they are used in the medical field, standardized methods for their disinfection must be implemented and closely followed. Also, manufacturers should become aware of the needs of the medical community regarding such devices. Thus, they might avoid building devices that—while being alluring for the medical

sector—do not respect the demands for hygiene required for medicinal products. However, the most efficient personal action one can take to avoid transmission of bacteria, viruses, and other pathogens remains the proper disinfection of the hands before and after every patient interaction—this is a fact independent of the kind of device or any operating system or stated purpose.

Future studies should also take the specific profession of the staff as well as their level in the hierarchy into account. Their attitude towards using the deBac-app—based procedure compared to regular hand hygiene using alcohol-based disinfection solutions should also be evaluated. Also, the expenditure of time for implementing the procedure and other cost implicating variables will need to be addressed as the gathered results would be important factors for decision makers.

Acknowledgments

The authors would like to thank Gisela Rademacher and the technical staff of the Institute for Medical Microbiology and Hospital Epidemiology at Hannover Medical School as well as Prof Dr Herbert Matthies (PLRI) for their effort, expertise, and contribution while conducting this study. We would also like to thank Dr. Marianne Behrends (PLRI) as well as Regina Schmeer, MSc (Nursing Department) for giving us the chance to evaluate the iPads that were deployed as mentioned in [17] with respect to hygiene. The study was financed solely based on institutional funds of the Hannover Medical School. We acknowledge support by Deutsche Forschungsgemeinschaft for the publication costs. The paper was screened against the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) checklist [36] to assess methodological weaknesses.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Standardized disinfection process of an iPad as guided by the corresponding application (“deBac-app”), the app documents the cleaning attempts of the device’s frame, front and back side.

[PDF File (Adobe PDF File), 132KB - [jmir_v15i8e176_app1.pdf](#)]

References

1. Rosenthal VD, Bijie H, Maki DG, Mehta Y, Apisarnthanarak A, Medeiros EA, INICC members. International Nosocomial Infection Control Consortium (INICC) report, data summary of 36 countries, for 2004-2009. *Am J Infect Control* 2012 Jun;40(5):396-407. [doi: [10.1016/j.ajic.2011.05.020](#)] [Medline: [21908073](#)]
2. Gagliotti C, Balode A, Baquero F, Degener J, Grundmann H, Gür D, EARS-Net Participants (Disease Specific Contact Points for AMR). Escherichia coli and Staphylococcus aureus: bad news and good news from the European Antimicrobial Resistance Surveillance Network (EARS-Net, formerly EARSS), 2002 to 2009. *Euro Surveill* 2011;16(11):pii/19834 [FREE Full text] [Medline: [21435327](#)]
3. Sehulster L, Chinn RY, CDC, HICPAC. Guidelines for environmental infection control in health-care facilities. Recommendations of CDC and the Healthcare Infection Control Practices Advisory Committee (HICPAC). *MMWR Recomm Rep* 2003 Jun 6;52(RR-10):1-42 [FREE Full text] [Medline: [12836624](#)]
4. Dasari KB, White SM, Pateman J. Survey of iPhone usage among anaesthetists in England. *Anaesthesia* 2011 Jul;66(7):630-631. [doi: [10.1111/j.1365-2044.2011.06747.x](#)] [Medline: [21682718](#)]
5. Payne KB, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Med Inform Decis Mak* 2012;12:121 [FREE Full text] [doi: [10.1186/1472-6947-12-121](#)] [Medline: [23110712](#)]
6. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. *J Med Syst* 2012 Oct;36(5):3135-3139. [doi: [10.1007/s10916-011-9798-7](#)] [Medline: [22052129](#)]
7. Gould D. Commentary: Ulger F et al. (2009). Are we aware how contaminated our mobile phones with nosocomial pathogens? *Nurs Crit Care* 2009;14(4):213-214. [doi: [10.1111/j.1478-5153.2009.00345.x](#)] [Medline: [19531040](#)]
8. Borer A, Gilad J, Smolyakov R, Eskira S, Peled N, Porat N, et al. Cell phones and Acinetobacter transmission. *Emerg Infect Dis* 2005 Jul;11(7):1160-1161 [FREE Full text] [doi: [10.3201/eid1107.050221](#)] [Medline: [16032803](#)]

9. Goldblatt JG, Krief I, Klonsky T, Haller D, Milloul V, Sixsmith DM, et al. Use of cellular telephones and transmission of pathogens by medical staff in New York and Israel. *Infect Control Hosp Epidemiol* 2007 Apr;28(4):500-503. [doi: [10.1086/513446](https://doi.org/10.1086/513446)] [Medline: [17385162](https://pubmed.ncbi.nlm.nih.gov/17385162/)]
10. Brady RR, Verran J, Damani NN, Gibb AP. Review of mobile communication devices as potential reservoirs of nosocomial pathogens. *J Hosp Infect* 2009 Apr;71(4):295-300. [doi: [10.1016/j.jhin.2008.12.009](https://doi.org/10.1016/j.jhin.2008.12.009)] [Medline: [19168261](https://pubmed.ncbi.nlm.nih.gov/19168261/)]
11. Ramesh J, Carter AO, Campbell MH, Gibbons N, Powlett C, Moseley H, et al. Use of mobile phones by medical staff at Queen Elizabeth Hospital, Barbados: evidence for both benefit and harm. *J Hosp Infect* 2008 Oct;70(2):160-165. [doi: [10.1016/j.jhin.2008.06.007](https://doi.org/10.1016/j.jhin.2008.06.007)] [Medline: [18701190](https://pubmed.ncbi.nlm.nih.gov/18701190/)]
12. Akinyemi KO, Atapu AD, Adetona OO, Coker AO. The potential role of mobile phones in the spread of bacterial infections. *J Infect Dev Ctries* 2009;3(8):628-632 [FREE Full text] [Medline: [19801807](https://pubmed.ncbi.nlm.nih.gov/19801807/)]
13. Datta P, Rani H, Chander J, Gupta V. Bacterial contamination of mobile phones of health care workers. *Indian J Med Microbiol* 2009;27(3):279-281 [FREE Full text] [doi: [10.4103/0255-0857.53222](https://doi.org/10.4103/0255-0857.53222)] [Medline: [19584520](https://pubmed.ncbi.nlm.nih.gov/19584520/)]
14. Ulger F, Esen S, Dilek A, Yanik K, Gunaydin M, Leblebicioglu H. Are we aware how contaminated our mobile phones with nosocomial pathogens? *Ann Clin Microbiol Antimicrob* 2009;8:7 [FREE Full text] [doi: [10.1186/1476-0711-8-7](https://doi.org/10.1186/1476-0711-8-7)] [Medline: [19267892](https://pubmed.ncbi.nlm.nih.gov/19267892/)]
15. Sadat-Ali M, Al-Omran AK, Azam Q, Bukari H, Al-Zahrani AJ, Al-Turki RA, et al. Bacterial flora on cell phones of health care providers in a teaching institution. *Am J Infect Control* 2010 Jun;38(5):404-405. [doi: [10.1016/j.ajic.2009.08.007](https://doi.org/10.1016/j.ajic.2009.08.007)] [Medline: [20363049](https://pubmed.ncbi.nlm.nih.gov/20363049/)]
16. Kramer A, Schwebke I, Kampf G. How long do nosocomial pathogens persist on inanimate surfaces? A systematic review. *BMC Infect Dis* 2006;6:130 [FREE Full text] [doi: [10.1186/1471-2334-6-130](https://doi.org/10.1186/1471-2334-6-130)] [Medline: [16914034](https://pubmed.ncbi.nlm.nih.gov/16914034/)]
17. Albrecht U, Behrends M, Schmeer R, Matthies HK, von Jan U. Usage of Multilingual Mobile Translation Applications in Clinical Settings. *JMIR Mhealth and Uhealth* 2013 Apr 23;1(1):e4. [doi: [10.2196/mhealth.2268](https://doi.org/10.2196/mhealth.2268)]
18. PLRI MedAppLab. deBac-app. 2013. URL: <https://itunes.apple.com/us/app/debac-app/id453976138?mt=8&ign-mpt=uo%3D2> [accessed 2013-02-08] [WebCite Cache ID 6EHfghOKx]
19. Gebert A, Preiss G. A simple method for the acquisition of high-quality digital images from analog scanning electron microscopes. *J Microsc* 1998 Sep;191(3):297-302. [Medline: [9767494](https://pubmed.ncbi.nlm.nih.gov/9767494/)]
20. Apple Inc. How to clean Apple products. 2013. URL: <http://support.apple.com/kb/ht3226> [accessed 2013-02-08] [WebCite Cache ID 6EHeciNaO]
21. Stenehjem E, Rimland D. MRSA nasal colonization burden and risk of MRSA infection. *Am J Infect Control* 2013 May;41(5):405-410. [doi: [10.1016/j.ajic.2012.07.017](https://doi.org/10.1016/j.ajic.2012.07.017)] [Medline: [23261345](https://pubmed.ncbi.nlm.nih.gov/23261345/)]
22. Panousi MN, Williams GJ, Girdlestone S, Hiom SJ, Maillard JY. Evaluation of alcohol wipes used during aseptic manufacturing. *Lett Appl Microbiol* 2009 May;48(5):648-651. [doi: [10.1111/j.1472-765X.2009.02574.x](https://doi.org/10.1111/j.1472-765X.2009.02574.x)] [Medline: [19228287](https://pubmed.ncbi.nlm.nih.gov/19228287/)]
23. Noether GE. Sample Size Determination for Some Common Nonparametric Tests. *Journal of the American Statistical Association* 1987 Jun;82(398):645 [FREE Full text]
24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979 Mar;86(2):420-428. [doi: [10.1037/0033-2909.86.2.420](https://doi.org/10.1037/0033-2909.86.2.420)] [Medline: [18839484](https://pubmed.ncbi.nlm.nih.gov/18839484/)]
25. Leggett MJ, McDonnell G, Denyer SP, Setlow P, Maillard JY. Bacterial spore structures and their protective role in biocide resistance. *J Appl Microbiol* 2012 Sep;113(3):485-498. [doi: [10.1111/j.1365-2672.2012.05336.x](https://doi.org/10.1111/j.1365-2672.2012.05336.x)] [Medline: [22574673](https://pubmed.ncbi.nlm.nih.gov/22574673/)]
26. Brady RR, Fraser SF, Dunlop MG, Paterson-Brown S, Gibb AP. Bacterial contamination of mobile communication devices in the operative environment. *J Hosp Infect* 2007 Aug;66(4):397-398. [doi: [10.1016/j.jhin.2007.04.015](https://doi.org/10.1016/j.jhin.2007.04.015)] [Medline: [17573157](https://pubmed.ncbi.nlm.nih.gov/17573157/)]
27. Brady RR, Wasson A, Stirling I, McAllister C, Damani NN. Is your phone bugged? The incidence of bacteria known to cause nosocomial infection on healthcare workers' mobile phones. *J Hosp Infect* 2006 Jan;62(1):123-125. [doi: [10.1016/j.jhin.2005.05.005](https://doi.org/10.1016/j.jhin.2005.05.005)] [Medline: [16099536](https://pubmed.ncbi.nlm.nih.gov/16099536/)]
28. Jeske HC, Tiefenthaler W, Hohlrieder M, Hinterberger G, Benzer A. Bacterial contamination of anaesthetists' hands by personal mobile phone and fixed phone use in the operating theatre. *Anaesthesia* 2007 Sep;62(9):904-906. [doi: [10.1111/j.1365-2044.2007.05172.x](https://doi.org/10.1111/j.1365-2044.2007.05172.x)] [Medline: [17697216](https://pubmed.ncbi.nlm.nih.gov/17697216/)]
29. Kilpatrick C, Pittet D. WHO SAVE LIVES: Clean Your Hands global annual campaign. A call for action: 5 May 2011. *Infection* 2011 Apr;39(2):93-95. [doi: [10.1007/s15010-011-0106-8](https://doi.org/10.1007/s15010-011-0106-8)] [Medline: [21484356](https://pubmed.ncbi.nlm.nih.gov/21484356/)]
30. Brady RR, Hunt AC, Visvanathan A, Rodrigues MA, Graham C, Rae C, et al. Mobile phone technology and hospitalized patients: a cross-sectional surveillance study of bacterial colonization, and patient opinions and behaviours. *Clin Microbiol Infect* 2011 Jun;17(6):830-835. [doi: [10.1111/j.1469-0691.2011.03493.x](https://doi.org/10.1111/j.1469-0691.2011.03493.x)] [Medline: [21615607](https://pubmed.ncbi.nlm.nih.gov/21615607/)]
31. Morioka I, Tabuchi Y, Takahashi Y, Oda Y, Nakai M, Yanase A, et al. [Bacterial contamination of mobile phones shared in hospital wards and the consciousness and behavior of nurses about biological cleanliness]. *Nihon Eiseigaku Zasshi* 2011 Jan;66(1):115-121 [FREE Full text] [Medline: [21358142](https://pubmed.ncbi.nlm.nih.gov/21358142/)]
32. Sumritvanicha A, Chintanavilas K, Apisarnthanarak A. Prevalence and type of microorganisms isolated from house staff's mobile phones before and after alcohol cleaning. *Infect Control Hosp Epidemiol* 2011 Jun;32(6):633-634. [doi: [10.1086/660204](https://doi.org/10.1086/660204)] [Medline: [21558783](https://pubmed.ncbi.nlm.nih.gov/21558783/)]

33. Hodges LR, Rose LJ, O'Connell H, Arduino MJ. National validation study of a swab protocol for the recovery of *Bacillus anthracis* spores from surfaces. *J Microbiol Methods* 2010 May;81(2):141-146. [doi: [10.1016/j.mimet.2010.02.010](https://doi.org/10.1016/j.mimet.2010.02.010)] [Medline: [20193714](https://pubmed.ncbi.nlm.nih.gov/20193714/)]
34. Moore G, Griffith C. Problems associated with traditional hygiene swabbing: the need for in-house standardization. *J Appl Microbiol* 2007 Oct;103(4):1090-1103. [doi: [10.1111/j.1365-2672.2007.03330.x](https://doi.org/10.1111/j.1365-2672.2007.03330.x)] [Medline: [17897214](https://pubmed.ncbi.nlm.nih.gov/17897214/)]
35. Galvin S, Dolan A, Cahill O, Daniels S, Humphreys H. Microbial monitoring of the hospital environment: why and how? *J Hosp Infect* 2012 Nov;82(3):143-151. [doi: [10.1016/j.jhin.2012.06.015](https://doi.org/10.1016/j.jhin.2012.06.015)] [Medline: [23022372](https://pubmed.ncbi.nlm.nih.gov/23022372/)]
36. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004 Mar;94(3):361-366. [Medline: [14998794](https://pubmed.ncbi.nlm.nih.gov/14998794/)]

Abbreviations

CFU: colony forming units

HCW: health care workers

IQR: interquartile range

MRSA: methicillin resistant *Staphylococcus aureus*

PDA: personal digital assistant

PLRI: Peter L Reichertz Institute for Medical Informatics

Edited by G Eysenbach; submitted 31.03.13; peer-reviewed by A Friedrich, B Baskerville; comments to author 25.04.13; revised version received 26.05.13; accepted 13.06.13; published 14.08.13.

Please cite as:

Albrecht UV, von Jan U, Sedlacek L, Groos S, Suerbaum S, Vonberg RP

Standardized, App-Based Disinfection of iPads in a Clinical and Nonclinical Setting: Comparative Analysis

J Med Internet Res 2013;15(8):e176

URL: <http://www.jmir.org/2013/8/e176/>

doi: [10.2196/jmir.2643](https://doi.org/10.2196/jmir.2643)

PMID: [23945468](https://pubmed.ncbi.nlm.nih.gov/23945468/)

©Urs-Vito Albrecht, Ute von Jan, Ludwig Sedlacek, Stephanie Groos, Sebastian Suerbaum, Ralf-Peter Vonberg. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 14.08.2013. 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

Efficacy of a Text Message-Based Smoking Cessation Intervention for Young People: A Cluster Randomized Controlled Trial

Severin Haug¹, PhD; Michael P Schaub¹, PhD; Vigeli Venzin², MA; Christian Meyer³, PhD; Ulrich John³, PhD

¹Swiss Research Institute for Public Health and Addiction, an associated Institute of the University of Zurich, Zurich, Switzerland

²Cantonal Office for Secondary Education Zurich, Zurich, Switzerland

³Institute of Social Medicine and Prevention, University Medicine Greifswald, Greifswald, Germany

Corresponding Author:

Severin Haug, PhD

Swiss Research Institute for Public Health and Addiction

an associated Institute of the University of Zurich

Konradstrasse 32

Zurich, PO Box 8031

Switzerland

Phone: 41 44 448 11 74

Fax: 41 44 448 11 70

Email: severin.haug@isgf.uzh.ch

Abstract

Background: Smoking prevalence remains high, particularly among adolescents and young adults with lower educational levels, posing a serious public health problem. There is limited evidence of effective smoking cessation interventions in this population.

Objective: To test the efficacy of an individually tailored, fully automated text messaging (short message service, SMS)-based intervention for smoking cessation in young people.

Methods: A 2-arm cluster randomized controlled trial, using school class as the randomization unit, was conducted to test the efficacy of the SMS text messaging intervention compared to an assessment-only control group. Students who smoked were proactively recruited via online screening in vocational school classes. Text messages, tailored to demographic and smoking-related variables, were sent to the participants of the intervention group at least 3 times per week over a period of 3 months. A follow-up assessment was performed 6 months after study inclusion. The primary outcome measure was 7-day smoking abstinence. Secondary outcomes were 4-week smoking abstinence, cigarette consumption, stage of change, and attempts to quit smoking. We used regression models controlling for baseline differences between the study groups to test the efficacy of the intervention. Both complete-case analyses (CCA) and intention-to-treat analyses (ITT) were performed. Subgroup analyses were conducted for occasional and daily smokers.

Results: A total of 2638 students in 178 vocational school classes in Switzerland participated in the online screening. Overall, 1012 persons met the inclusion criteria for study participation, and 755 persons (74.6%) participated in the study (intervention: n=372; control: n=383). Of the 372 program participants, 9 (2.4%) unsubscribed from the program during the intervention period. Six-month follow-up data were obtained for 559 study participants (74.0%). The 7-day smoking abstinence rate at follow-up was 12.5% in the intervention group and 9.6% in the control group (ITT: $P=.92$). No differences between the study groups were observed in 4-week point prevalence abstinence rates. The decrease in the mean number of cigarettes smoked per day from baseline to follow-up was higher in the intervention group than in the control group (ITT: $P=.002$). No differences between the groups were observed in stage of change (ITT: $P=.82$) and quit attempts (ITT: $P=.38$). The subgroup analyses revealed lower cigarette consumption in both occasional and daily smokers in the intervention group compared to the control group. Occasional smokers in the intervention group made more attempts to quit smoking than occasional smokers in the control group.

Conclusions: This study demonstrated the potential of an SMS text message-based intervention to reach a high proportion of young smokers with low education levels. The intervention did not have statistically significant short-term effects on smoking cessation; however, it resulted in statistically significant lower cigarette consumption. Additionally, it resulted in statistically significant more attempts to quit smoking in occasional smokers.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 19739792; <http://www.controlled-trials.com/ISRCTN19739792> (Archived by WebCite at <http://webcitation.org/6IGETTHmr>).

(*J Med Internet Res* 2013;15(8):e171) doi:[10.2196/jmir.2636](https://doi.org/10.2196/jmir.2636)

KEYWORDS

smoking cessation; text messaging (SMS); young people; school; students

Introduction

Tobacco use is a major cause of the global disease burden and is the single most preventable cause of death in the world [1]. A survey of 15- and 16-year-old adolescents covering 36 European countries revealed that the smoking prevalence rate of 28% having used cigarettes during the past 30 days has remained stable over the past 4 years [2]. Smoking continues to be a serious problem, particularly in adolescents and young adults with lower education levels [3].

There is limited evidence of smoking cessation interventions demonstrating efficacy in young people [4,5]. The 2006 Cochrane Review for smoking cessation interventions for those younger than 20 years identified only 15 trials of sufficient quality, of which only 1 [6] found statistically significant evidence of an intervention effect [4]. The authors acknowledged that there is a need for well-designed, adequately powered trials of cessation interventions. The authors concluded that complex approaches, including elements sensitive to stage of change, achieved moderate long-term success, whereas the efficacy of psychosocial and pharmacological interventions could not be demonstrated. A more recent but less systematic review from 2008 [5] suggested that delivering smoking cessation programs for youth in contexts that are geared to youth, interventions addressing cognitive behavioral, motivational and social influence contents, and programs with at least 5 sessions were most effective. Since the publication of these reviews, additional randomized controlled trials of adolescent smoking interventions have been reported, from which 2 found a treatment effect at 6-month follow-up: Pbert et al [7] provided brief counseling by the pediatric provider followed by 1 visit and 4 telephone calls by older peer counselors; Peterson et al [8] provided proactive telephone counseling of high school juniors.

Beyond intervention effectiveness, intervention reach and retention are major challenges of smoking cessation interventions in young people [9,10]. Reaching a large proportion of adolescent smokers has been difficult. Less than 50% of smokers are typically recruited in school-based smoking cessation programs [4,11]. However, a large reach is essential for the efficacy of an intervention at the population level. For a large reach, proactive recruitment strategies are needed that address all persons among a given target population. All smokers should receive the invitation to take part in smoking cessation. Such recruitment, in combination with low-threshold interventions, seems promising [7,8].

Mobile phone text messaging (short message service, SMS) is very popular among adolescents and young adults and has the potential to deliver smoking cessation support to large proportions of the population. Among 12- to-19-year-old

adolescents from Switzerland, 98% owned a mobile phone in 2010; use of the mobile phone was the most frequent leisure time activity in this population group [12]. Reading and sending SMS text messages were the most frequent activities when using a mobile phone [12].

By using expert system technology that provides information based on individual demographic- or smoking-related characteristics, electronic communication technology can be a viable time- and cost-saving alternative to interpersonal counseling [13]. SMS text messaging provides an opportunity for individualized and interactive information delivery that may easily be accessed, independent of time and place. A recent Cochrane Review including 5 randomized or quasi-randomized studies revealed an overall long-term benefit of mobile phone interventions for smoking cessation in adults [14], although there was a high level of statistical heterogeneity in the pooled results. A large, methodologically sound trial was conducted in Great Britain to test the efficacy of SMS text message-based smoking cessation interventions in adults motivated to quit smoking [15]. Within this study, smokers who intended to quit within the subsequent month received motivational messages and behavioral-change support over a period of 26 weeks. The messages were matched to participants' demographic and smoking-related characteristics gathered at baseline. Additionally, participants could request instant messages aimed at craving or lapse situations. The program significantly improved smoking cessation rates at 6 months compared to a control group that received text messages unrelated to quitting (9% vs 4%, respectively).

To date, neither randomized controlled trials testing the efficacy of smoking cessation interventions employing SMS text messaging in adolescents and young adults nor trials testing the efficacy of SMS text message interventions in proactively recruited smokers have been reported. In 3 pilot studies in which young adult smokers, irrespective of their motivation to quit, were proactively invited to an SMS text message-based smoking cessation intervention, high participation and retention proportions were achieved [16-18].

Within the present cluster randomized trial, we tested the efficacy of an SMS text message-based intervention for smoking cessation in a sample of proactively recruited students with varying motivation to quit. Vocational school students were chosen as the target population because smoking prevalence rates in this subgroup of adolescents and young adults with heterogeneous educational levels are high [3].

Methods

Study Design

A 2-arm cluster randomized controlled trial (ISRCTN: 19739792 assigned on May 20, 2011) was conducted to test the efficacy of the program SMS-COACH, an SMS text message-based intervention for smoking cessation in adolescents and young adults, compared to an assessment-only control group. The trial was undertaken in Switzerland, and participants were recruited between October 2011 and May 2012. The 6-month follow-ups were conducted between April and December 2012; the study protocol was published on January 19, 2012 [19]. Students in vocational schools were proactively invited to participate, irrespective of their intention to quit. The smoking cessation text messages were primarily based on the Health Action Process Approach (HAPA) [20] and included cognitive behavioral and motivational components according to this model. Text messages were sent to the participants over a period of 3 months and were tailored according to data gathered at baseline and a weekly SMS text message assessment. At the 6-month follow-up, we expected a higher 7-day point prevalence smoking abstinence rate in students in the intervention group compared to students in the assessment-only control group. Secondary outcome measures were 4-week point prevalence smoking abstinence, the number of cigarettes smoked per day, stage of change, and number of attempts to quit smoking. The study protocol was approved by the Local Ethics Committee of the Canton of Zurich, Switzerland (date of approval: March 15, 2011; No: KEK-StV-Nr. 05/11). The trial was executed in compliance with the Helsinki Declaration.

The study was implemented as described in the study protocol [19] with the following modifications: (1) because of smaller class sizes than expected and time restrictions, we could not reach the targeted sample size of 910 study participants, but enrolled 755 study participants; (2) self-efficacy for smoking cessation could not be assessed at follow-up and used as a secondary outcome measure because the rating scale to assess this variable [20] could not be applied in the telephone interviews conducted at follow-up; and (3) nicotine dependence could not be calculated for occasional smokers using the Heaviness of Smoking Index [21]. Therefore, we used number of cigarettes smoked per day as an indicator of nicotine dependence and as an outcome variable.

Participant Recruitment and Baseline Assessment

Smoking students were recruited at vocational schools in Switzerland. Contact teachers for addiction prevention or headmasters of 57 vocational schools in German-speaking regions of Switzerland were invited to enroll some of their classes in a study testing the efficacy of an SMS text message-based smoking cessation program. Teachers from the 24 participating vocational schools scheduled 1 school hour per class for screening of eligibility criteria, study information, baseline assessments, and program registration. Study participants were recruited by study assistants (graduate students of psychology). The study assistants invited all students from a school class to participate in an online health survey during a regular school lesson reserved for health education. They

informed the students that some people would be invited to participate in a study testing the efficacy of an SMS text message intervention for health promotion. To decrease reporting bias, the study assistants did not provide more information about the purpose of the study before the screening of eligibility criteria was completed.

Afterwards, the students were invited to complete an online screening. The screening included the assessment of demographic data, alcohol consumption, weekly physical activity, smoking status, and ownership of a mobile phone. Inclusion criteria for study participation were (1) daily or occasional cigarette smoking (at least 4 cigarettes in the preceding month and at least 1 cigarette during the preceding week), and (2) ownership of a mobile phone. Subsequently, eligible persons were informed by the online program about the aim of the study, the intervention arms, assessments, reimbursement, and data protection. Study information was provided online and in paper form by the study assistants. The equivalent of €8 was offered as reimbursement to all study participants for participation at the 6-month follow-up assessment. Additionally, the equivalent of €0.80 was offered as reimbursement to the participants of the intervention group for each SMS text message response to the weekly SMS text message assessments in the program. After receiving informed consent online, all study participants were invited to choose a username and to provide their mobile phone number. Subsequently, the following smoking-related variables were assessed: stage of change, number of cigarettes smoked per day, past quit attempts, and age of smoking onset. Afterwards, study participants of the intervention group received further information about the operation of the program. Control group participants were informed that they were assigned to the control group and could not participate in the SMS text message program.

Randomization and Allocation Concealment

To avoid spillover effects within school classes, we used cluster randomization with school class as the randomization unit. Because of the heterogeneity of students in the different vocational schools (ie, gender or course of study), we used separate randomization lists for each vocational school (stratified randomization). Furthermore, to approximate equality of sample sizes in the study groups, we used block randomization with computer-generated, randomly permuted blocks of 4 cases [22].

The study assistants who conducted the baseline assessment in the vocational schools were blinded concerning group allocation for each of the school classes. Additionally, group allocation was not released to study participants until they provided informed consent, username, mobile phone number, and baseline data for the smoking-related variables. The study assistants who conducted the computer-assisted telephone interviews at follow-up were blinded to group allocation when assessing the primary and secondary outcome measures.

Sample Size Calculation

Based on results of a study that tested the efficacy of telephone counseling for smoking cessation in high school students [8], we expected an 8% difference in 7-day point prevalence

abstinence rates between the intervention and the control condition at 6-month follow-up assessment (25% vs 17%, respectively). To achieve a power of .80 with a significance level of .05 using a chi-square test (χ^2), a sample size of $n=406$ in each study group was necessary. Because students were nested within school classes, we also needed to consider a potential design effect of 1.12 (average cluster size $n=7$; intraclass correlation coefficient: 0.02), which resulted in a required sample size of $n=455$ per study group.

Intervention

Technological Background

The text messaging intervention, SMS-COACH, was fully automated and based on Internet technology using a Linux, Apache, MySQL, and PHP (LAMP) system. The program used in the present study was an extended and modified version of a previous version that had been tested successfully in pilot studies [16-18]. All incoming and outgoing text messages were automatically recorded. Incoming messages were analyzed immediately.

Theoretical Background

The program was primarily based on the HAPA [20]. This health behavior model suggests a distinction between motivation processes resulting in goal setting and volition processes leading to the actual health behavior. The approach combines 3 nonactive stages (precontemplation, contemplation, and preparation) and 2 active stages (action and maintenance). Within the initial 2 stages, outcome expectancies, risk perception, and perceived self-efficacy are important social-cognitive predictors to develop an intention to act. Within the subsequent intentional stage (preparation), planning processes are crucial to achieve the desired action. Once an action has been initiated, self-regulatory skills are important to maintain the healthy behavior. In addition to the HAPA, we used intervention elements derived from the Social Norms Approach [23] and implementation intentions, which are if-then plans that link situational cues with responses that are effective in attaining a desired outcome [24].

Intervention Elements

The intervention program consisted of (1) an online assessment of individual smoking behavior and attitudes toward smoking cessation, (2) a weekly SMS text message assessment of smoking-related target behaviors, (3) 2 weekly text messages tailored to the data of the online and the SMS text message assessments, and (4) an integrated quit day preparation and relapse-prevention program.

Online Baseline Assessment

In addition to the screening questions and the previously mentioned smoking-related variables that were assessed in both study groups at the baseline assessment, participants of the intervention group received online questions assessing (1) outcome expectancies of smoking cessation, (2) situations or circumstances in which craving for cigarettes usually occurs, (3) alternative strategies to handle these craving situations, and (4) costs per cigarette package.

Weekly Text Message Assessment

During the 3-month intervention period, participants in the intervention group received 1 text message per week to assess smoking-related target behavior. This question could be answered easily by typing a single letter or number, using the reply function of the mobile phone. The weekly SMS text message assessment question was sent at a fixed time point each week (6 pm on the weekday of study registration). The content of the question depended on the HAPA stage as well as on the number of the intervention week.

For all participants, the HAPA stage was assessed in even weeks by the question: "Have you recently smoked cigarettes?" with the following response options (1) "Yes, and I do not intend to quit" (precontemplation), (2) "Yes, but I am considering quitting" (contemplation), (3) "Yes, but I seriously intend to quit" (preparation), or (4) "No, I quit smoking" (action). This question assessed both smoking status and intention to quit over time. The responses to this question allowed the tailoring of the SMS text message feedback according to the current HAPA stage [25].

In odd weeks, we assessed the number of cigarettes smoked per day or week (depending on smoking status: daily/occasionally) in smokers in the preintentional stages (precontemplation and contemplation). We also assessed whether smokers in the intention or action stage applied the individually chosen strategies to cope with craving situations (eg, "Did you apply the following strategy recently? When I am at a party, I distract myself from smoking by dancing.").

Individually Tailored Text Messages

At the first level, text messages were tailored to the HAPA stage. Persons in the preintentional stages received text messages addressing (1) the risks of smoking, (2) the monetary costs of smoking, (3) the social norms of smoking, (4) outcome expectancies, and (5) motivation to reduce the number of cigarettes smoked per day (daily smokers) or week (occasional smokers). Persons in the intentional stage received text messages that (1) motivated them to use social support for smoking cessation, (2) provided strategies to cope with craving situations, and (3) provided tips for preparing for smoking cessation (eg, reducing the number of cigarettes, identifying craving situations). Persons in the action stage received text messages (1) motivating them to reward themselves for staying abstinent, (2) providing strategies to cope with craving situations, and (3) motivating them to use social support for staying abstinent.

On the second level, the text messages were tailored according to the individual information provided at the baseline assessment as well as through the weekly SMS text message assessments. Examples of text messages are displayed in the study protocol of this trial [19] or in [Multimedia Appendix 1](#).

Integrated Program for Quit Day Preparation and Relapse Prevention

Persons in the preparation and action stage had the possibility to additionally participate in an integrated program for quit day preparation and relapse prevention. Program participants in these stages were informed biweekly about this option. After

entering a scheduled quit date, the program provided up to 2 daily text messages (weeks -1 to +1: 2 daily SMS text messages; weeks +2 and +3: 1 daily text message) to prepare for the quit day and to prevent relapse afterwards.

Number of Text Messages Sent to the Participants

Participants who did not use the integrated program for quit day preparation and relapse prevention received a total of 37 text messages (1 welcome message, 11 assessment messages, 24 tailored feedback messages, 1 goodbye message). Participants, who used the quit day preparation and relapse-prevention program for the whole period from 1 week before the scheduled quit date until 3 weeks afterwards, received an additional 42 text messages.

Control Group

Study participants in the assessment-only control group did not receive any of the previously described intervention elements of the SMS-COACH program.

Baseline Measures

The screening assessment included the following demographic variables: gender, age, school education, and immigration background. Common Swiss levels of educational attainment were assessed: (1) none, (2) secondary school, (3) extended secondary school, and (4) technical or high school. We assessed the country of birth of both parents of the students to identify a potential immigrant background. Based on this information, participants were assigned to one of the following categories: (1) neither parent born outside Switzerland, (2) 1 parent born outside Switzerland, or (3) both parents born outside Switzerland.

The following health-related variables were assessed: physical activity and alcohol use. Self-reported moderate to vigorous physical activity was measured by a question derived from the Health Behavior in School-Aged Children (HBSC) study [26]: "Outside school, how many hours a week do you exercise or participate in sports that make you sweat or out of breath?" Alcohol consumption was assessed using the first 3 items about consumption of the Alcohol Use Disorder Identification Test (AUDIT-C), [27,28]. The AUDIT-C assesses drinking quantity, drinking frequency, and binge drinking. Based on recent recommendations [29], we used the gender-specific cut-off values for the AUDIT-C total score, ≥ 4 for men and ≥ 3 for women, to determine whether hazardous drinking was present.

Tobacco smoking was assessed using the question, "Are you currently smoking cigarettes or did you smoke in the past?" with the following response options: (1) I smoke cigarettes daily; (2) I smoke cigarettes occasionally, but not daily; (3) I smoked cigarettes in the past, but I do not smoke anymore; and (4) I have never smoked cigarettes or have smoked less than 100 cigarettes in my life. In occasional smokers, we additionally assessed the number of days they typically smoked per month and the total number of cigarettes smoked within the previous 7 days. In daily smokers and occasional smokers who smoked at least 4 cigarettes in the preceding month and at least 1 cigarette during the preceding week, we additionally assessed the following smoking-related variables: mean number of

cigarettes smoked per day, stage of change according to the HAPA, and number of previous quit attempts.

In daily smokers, we assessed the number of cigarettes smoked on a typical day. In occasional smokers, we initially assessed the typical number of smoking days per month; subsequently, the number of cigarettes smoked on a typical smoking day was assessed. For occasional smokers, the number of cigarettes smoked per day was computed by multiplying the typical number of smoking days per month by the number of cigarettes smoked on a typical smoking day divided by 30. The stage of change based on the HAPA was assessed by the following question: "Have you recently smoked cigarettes?" with the following response options (1) "Yes, and I do not intend to quit" (precontemplation), (2) "Yes, but I am considering quitting" (contemplation), and (3) "Yes, but I seriously intend to quit" (preparation). Previous quit attempts were assessed by the question: "Have you ever made a serious attempt to quit smoking?" with the response options (1) no, (2) yes, once, and (3) yes, more than once. Furthermore, we assessed age at smoking onset by the question: "How old were you when you started smoking periodically?"

Program Participation and Program Use

To evaluate acceptance of the program, we analyzed log files of the SMS text message system in which the number and content of incoming and outgoing text messages were recorded. The number of responses to the weekly SMS text message assessments and the number of program participants who unsubscribed from the program (program attrition) were examined. At follow-up, we also assessed usage of the SMS text messages by asking the participants whether they (1) read the SMS text message feedback messages thoroughly, (2) took only a short look at the feedback messages, or (3) did not read the feedback messages.

Follow-Up Measures

Computer-assisted telephone interviews were conducted at the 6-month follow-up assessment by trained interviewers. The following outcome variables were assessed during this interview: (1) smoking status, (2) 7-day smoking abstinence, (3) 4-week smoking abstinence, (4) mean number of cigarettes smoked per day, (5) stage of change according to the HAPA, and (6) quit attempts within the past 6 months preceding the follow-up. The main outcome criterion was 7-day point prevalence smoking abstinence.

For assessment of smoking status, the participants could indicate whether they smoked (1) daily, (2) occasionally, or (3) do not smoke anymore. Furthermore, 7-day point prevalence smoking abstinence (ie, not having smoked a puff within the past 7 days preceding the follow-up [23]), and 4-week point prevalence smoking abstinence were assessed. Among daily smokers, we assessed the number of cigarettes smoked on a typical day. Among occasional smokers, we initially assessed the typical number of smoking days per month and subsequently the number of cigarettes smoked on a typical smoking day. For occasional smokers, the number of cigarettes smoked per day was computed by multiplying the typical number of smoking days per month by the number of cigarettes smoked on a typical

smoking day divided by 30. In participants who indicated that they did not smoke anymore, the value for the number of cigarettes smoked per day was set to zero.

The HAPA stage was assessed by a similar question as at baseline. Participants indicating that they did not smoke anymore were assigned to the action stage. Quit attempts within the previous 6 months were assessed by the yes/no question: "Have you made a serious attempt to quit smoking within the previous 6 months?" For participants who indicated that they did not smoke anymore, a serious quit attempt was assumed.

Data Analysis

The data were analyzed using STATA software, version 10. To test for baseline equivalence of intervention and control individuals, chi-square tests for categorical variables and *t* tests for continuous variables were used. For the attrition analysis (study participants lost to follow-up), we also used chi-square tests for categorical variables and *t* tests for continuous variables. Baseline equivalence and lack of attrition bias were assumed for tests with $P > .10$.

We used regression models to verify the efficacy of the intervention on the different outcome measures. Logistic regression models were applied for the binary outcome variables (7-day and 4-week point prevalence smoking abstinence), negative binomial regression models were applied for the count data (number of cigarettes smoked per day), ordinal logistic regression models were used for ordinal data (stage of change), and multinomial logistic regression models were used for categorical outcomes (smoking status). To control for baseline differences, we additionally added the respective baseline variables as covariates to the regression models.

We conducted both complete-case analyses (CCA) considering all study participants with available follow-up data, and intention-to-treat (ITT) analyses. For the ITT analyses, we applied the multiple imputations procedure (MICE) of STATA, which imputed missing follow-up data by using all available baseline variables (demographic, health- and smoking-related variables). We created 30 imputed datasets. Given the clustered nature of the data (students within school classes), we computed robust variance estimators for all regression models using the *svy* command of STATA.

Because of significant baseline differences between the study groups, particularly in the percentage of occasional and daily smokers, and significant interaction effects of study condition \times smoking status for the number of cigarettes smoked per day ($P = .01$) and quit attempts within the previous 6 months ($P = .02$)

outcomes, we additionally conducted outcome analyses separately for occasional and daily smokers.

Results

Study Participation

Figure 1 presents a flowchart of the study participants. At the time of the online screening assessment in 178 school classes, a total of 2657 students were present. Among them, 2638 (99.3%) agreed to participate. Of these, 1012 persons met the inclusion criteria for study participation and 755 persons (74.6%) participated in the study. Ninety classes consisting of 372 students were randomly assigned to the intervention group and 88 classes consisting of 383 students were assigned to the control group. Follow-up assessments were completed in 287 (77.2%) study participants in the intervention group and 272 (71.0%) study participants in the control group.

Sample Characteristics

Baseline characteristics for the study sample are shown in Table 1.

Baseline differences between intervention and control group participants were found for the following variables: gender ($\chi^2_1 = 3.1, P = .08$), hazardous drinking ($\chi^2_1 = 4.8, P = .03$), smoking status ($\chi^2_1 = 13.3, P < .001$), number of cigarettes smoked per day ($t_{753} = 3.6, P < .001$), and age of onset of smoking ($t_{753} = -2.8, P = .005$).

We conducted ancillary separate analyses for occasional and daily smokers, and then we checked for baseline differences within these subgroups. Within the sample of occasional smokers, the following baseline differences between intervention and control group participants were found: (1) a higher percentage of male participants in the intervention group ($\chi^2_1 = 4.3, P = .04$), and (2) a higher number of cigarettes smoked per day in the intervention group ($t_{176} = -1.7, P = .09$). Within the sample of daily smokers, the following baseline differences between intervention and control group participants were found: (1) a lower percentage of hazardous drinking in the intervention group ($\chi^2_1 = 5.3, P = .02$), (2) lower cigarette consumption in the intervention group ($t_{575} = 1.9, P = .06$), and (3) a higher age of onset of smoking in the intervention group ($t_{575} = -1.8, P = .07$).

The attrition analysis revealed that individuals lost to follow-up were more likely to be daily smokers (81.1% vs 74.8%; $\chi^2_1 = 3.2, P = .07$) and smoked a higher number of cigarettes per day (11.5 vs 10.3; $t_{753} = 2.0, P = .048$).

Figure 1. Flowchart of study participants.

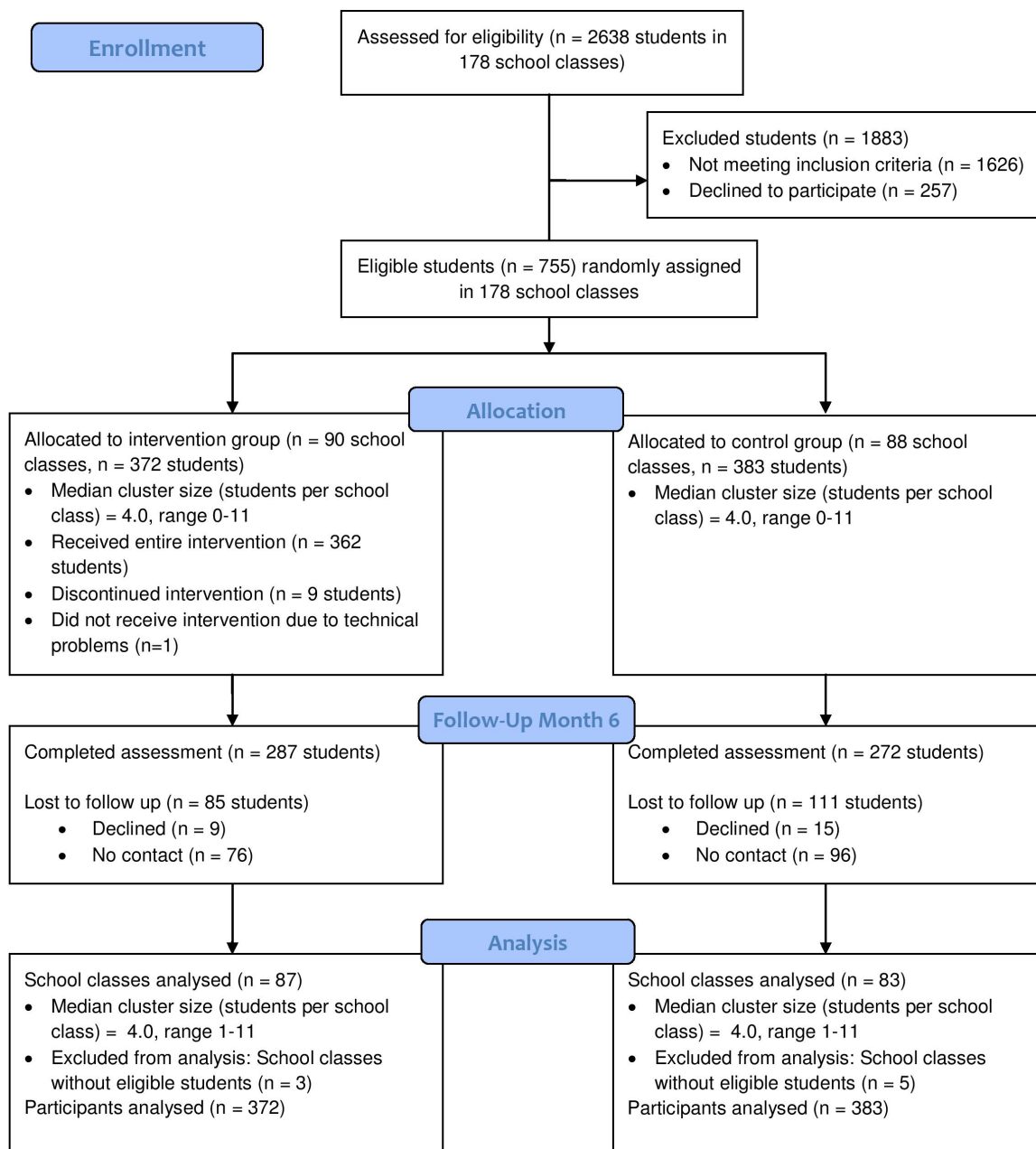


Table 1. Demographics and health- and smoking-related baseline characteristics of the study sample.

Variable	Intervention n=372	Control n=383	Total N=755	P ^a
Gender, n (%)				
Male	191 (51.3)	172 (44.9)	363 (48.1)	.08 ^b
Female	181 (48.7)	211 (55.1)	392 (51.9)	
Age, mean (SD)	18.2 (2.4)	18.3 (2.2)	18.2 (2.3)	.56 ^c
Immigration background, n (%)				
No immigration background	196 (52.7)	206 (53.8)	402 (53.2)	.89 ^b
One parent born outside Switzerland	75 (20.2)	79 (20.6)	154 (20.4)	
Both parents born outside Switzerland	101 (27.2)	98 (25.6)	199 (26.4)	
Education, n (%)				
None	14 (3.8)	11 (2.9)	25 (3.3)	.68 ^b
Secondary school	290 (78.0)	301 (78.6)	591 (78.3)	
Extended secondary school	59 (15.9)	56 (14.6)	115 (15.2)	
Technical or high school	9 (2.4)	15 (3.9)	24 (3.2)	
Hazardous drinking, n (%)				
No	76 (20.4)	55 (14.4)	131 (17.4)	.03 ^b
Yes	296 (79.6)	328 (85.6)	624 (82.6)	
Hours of extracurricular moderate to vigorous physical activity per week, mean (SD)	4.0 (4.6)	3.7 (5.0)	3.8 (4.8)	.37 ^c
Tobacco smoking status, n (%)				
Occasional smoker	109 (29.3)	69 (18.0)	178 (23.6)	<.001 ^b
Daily smoker	263 (70.7)	314 (82.0)	577 (76.4)	
Stage of change, n (%)				
No intention to quit	86 (23.1)	112 (29.2)	198 (26.2)	.11 ^b
Considering quitting	214 (57.5)	211 (55.1)	425 (56.3)	
Serious intention to quit	72 (19.4)	60 (15.7)	132 (17.5)	
Number of cigarettes smoked per day, mean (SD)	9.6 (7.2)	11.6 (7.9)	10.6 (7.6)	<.001 ^c
Age of onset of smoking, mean (SD)	15.1 (1.6)	14.8 (1.7)	15.0 (1.6)	.005 ^c
Previous quit attempts, n (%)				
0	148 (39.8)	141 (36.8)	289 (38.3)	.32 ^b
1	153 (41.1)	178 (46.5)	331 (43.8)	
2 or more	71 (19.1)	64 (16.7)	135 (17.9)	

^aP values for the comparison of intervention and control group participants.

^b χ^2 test.

^ct test.

Program Attrition and Program Use

During the program, which lasted for 3 months, 9 (2.4%) of the 372 participants in the intervention group unsubscribed from the program.

The mean number of replies to the weekly SMS text message assessments was 6.5 (SD 3.7). No reply was sent by 34

participants (9.1%), and all 11 replies were sent by 55 participants (14.8%).

Out of the 287 participants with valid follow-up data, 271 (94.4%) indicated that they regularly read the SMS text messages. Of these, 204 (75.3%) indicated that they read the SMS text messages thoroughly, whereas 67 participants (24.7%) reported that they took a short look at the feedback messages.

Program Efficacy

Smoking Abstinence

Table 2 presents 7-day and 4-week point prevalence smoking abstinence rates at follow-up for both study groups based on complete case data. Using CCA and ITT, the logistic regression analyses controlling for differences in baseline characteristics did not reveal any differences in 7-day or 4-week smoking abstinence rates at follow-up between the study groups for the total sample, the subgroup of occasional smokers and the subgroup of daily smokers.

Cigarette Consumption

Table 3 presents the mean number of cigarettes smoked per day at follow-up for both study groups based on complete case data. Both CCA and ITT revealed lower cigarette consumption in the intervention group than in the control group. Within baseline occasional smokers and baseline daily smokers, both CCA and ITT revealed lower cigarette consumption in the intervention group than in the control group.

Stage of Change

Table 4 presents the stage of change at follow-up for participants in both study groups. Using CCA and ITT, the regression models did not reveal differences in stages of change between the study groups for the total sample, the subgroup of occasional smokers, and the subgroup of daily smokers.

Quit Attempts

Based on complete case data of the total sample, 98 (36.3%) of 270 participants in the control group and 125 (43.7%) of 286 participants in the intervention group indicated that they made a quit attempt within the 6 months preceding follow-up (CCA: OR 1.17, 95% CI 0.81-1.71, $P=.40$; ITT: OR 1.18, 95% CI 0.81-1.72, $P=.38$). In baseline occasional smokers, 12 (43.1%) of 51 participants in the control group and 62 (68.9%) of 90 participants in the intervention group indicated a quit attempt (CCA: OR 2.79, 95% CI 1.36-5.73, $P=.006$; ITT: OR 2.48, 95% CI 1.24-4.93, $P=.01$). Using the subgroup of baseline daily smokers, 76 (34.7%) of 219 participants in the control group and 63 (32.1%) of 196 participants in the intervention group indicated a quit attempt (CCA: OR 0.87, 95% CI 0.55-1.37, $P=.54$; ITT: 0.95, 95% CI 0.62-1.46, $P=.82$).

Table 2. Point prevalence smoking abstinence rates at follow-up (complete-case data) and results of logistic regression analyses comparing abstinence rates in the study groups using complete-case analyses (CCA) and intention-to-treat analyses (ITT).

Sample	Control n (%)	Intervention n (%)	OR (95% CI)		P	
			CCA	ITT	CCA	ITT
Total sample^a						
7-day abstinence	26 (9.6)	36 (12.5)	1.02 (0.60-1.76)	1.03 (0.59-1.79)	.93	.92
4-week abstinence	15 (5.5)	18 (6.3)	0.87 (0.45-1.71)	0.97 (0.50-1.90)	.69	.92
Baseline occasional smokers^b						
7-day abstinence	10 (19.6)	25 (27.8)	1.64 (0.65-4.10)	1.56 (0.65-3.75)	.29	.32
4-week abstinence	4 (7.8)	13 (14.4)	1.99 (0.58-6.80)	2.06 (0.63-6.78)	.27	.23
Baseline daily smokers^c						
7-day abstinence	16 (7.2)	11 (5.6)	0.83 (0.35-1.92)	0.81 (0.36-1.81)	.65	.61
4-week abstinence	11 (5.0)	5 (2.5)	0.48 (0.13-1.80)	0.55 (0.17-1.77)	.27	.32

^aBased on 272 participants in the control group and 287 in the intervention group.

^bBased on 51 participants in the control group and 90 in the intervention group.

^cBased on 221 participants in the control group and 197 in the intervention group.

Table 3. Mean number of cigarettes smoked per day at follow-up (complete case data) and results of logistic regression analyses comparing cigarette consumption in the study groups using complete-case analyses (CCA) and intention-to-treat analyses (ITT).

Sample	Control mean (SD)	Intervention mean (SD)	t (df)	P		
				CCA	ITT	CCA
Total sample	10.0 (7.9)	7.5 (7.2)	-2.80 (164)	-3.18 (43.1)	.006	.002
Baseline occasional smokers	2.7 (3.2)	1.7 (2.4)	-2.32 (89)	-2.41 (125.9)	.02	.02
Baseline daily smokers	11.7 (7.7)	10.2 (7.1)	-2.22 (151)	-2.53 (37.5)	.03	.01

Table 4. Stage of change at follow-up based on complete case data and results of ordinal regression analyses comparing stage of change between the study groups using complete-case analyses (CCA) and intention-to-treat analyses (ITT).

Sample	Control n (%)	Intervention n (%)	<i>t</i> (<i>df</i>)		<i>P</i>	
			CCA	ITT	CCA	ITT
Total sample^a			0.33 (164)	0.23 (78.7)	.74	.82
Precontemplation	72 (26.7)	65 (22.7)				
Contemplation	133 (49.3)	134 (46.9)				
Preparation	29 (10.7)	32 (11.2)				
Action	36 (13.3)	55 (19.2)				
Baseline occasional smokers^b			1.60 (89)	1.33 (78.9)	.11	.18
Precontemplation	8 (15.7)	5 (5.6)				
Contemplation	22 (43.1)	38 (42.2)				
Preparation	6 (11.8)	6 (6.7)				
Action	15 (29.4)	41 (45.6)				
Baseline daily smokers^c			-0.24 (151)	-0.40 (89.4)	.81	.69
Precontemplation	64 (29.2)	60 (30.6)				
Contemplation	111 (50.7)	96 (49.0)				
Preparation	23 (10.5)	26 (13.3)				
Action	21 (9.6)	14 (7.1)				

^aBased on 270 participants in the control group and 286 in the intervention group.

^bBased on 51 participants in the control group and 90 participants in the intervention group.

^cBased on 219 participants in the control group and 196 participants in the intervention group.

Discussion

The study aimed to test the efficacy of an SMS text message–based intervention for smoking cessation in a sample of proactively recruited vocational school students with different motivation to quit. The study revealed 4 main findings: (1) a large percentage of smoking students participated in the program, (2) program attrition was low, (3) program participation resulted in lower cigarette consumption, but (4) no short-term effect of the intervention on smoking abstinence rates was found.

The proactive invitation for program participation in combination with the offer of a low-threshold intervention using SMS text messages allowed us to reach 3 of 4 smoking students (75%) for participation in the SMS-COACH program. Taking into account that 83% of the program participants were in the precontemplation or contemplation stage at baseline (ie, indicated no serious intention to quit), this high participation rate is of special relevance. Other school-based smoking cessation interventions conducted in German-speaking countries showed much lower participation rates of 37% [30] and 19% [11]. In-line with other recently developed smoking cessation approaches in adolescents [7,8], our results underscore the importance of proactive recruitment strategies and low-threshold interventions to attain a high participation rate. The flexibility of SMS text messaging to send and receive messages at any time, place, or setting, as well as the possibility to receive individually tailored information, might be responsible for the

high use and retention rates identified in this study. Nearly all program participants (98%) stayed logged in until the end of the 3-month program. The SMS text messages were read by almost all program participants (94%) and 9 of 10 program participants (91%) replied to the SMS text message assessments.

The finding that the intervention program resulted in lower cigarette consumption indicates that the intervention might promote smoking abstinence. The number of cigarettes smoked per day, which is closely related to nicotine dependence [31], has proved to be among the best predictors of smoking cessation in both adolescents and adults [32–34]. However, the main study outcome was 7-day point prevalence smoking abstinence assessed at the 6-month follow-up. This abstinence rate was 12.5% in the intervention group and 9.6% in the control group. After controlling for baseline differences, no significant intervention effect was found for this criterion. The separate subgroup analyses for daily and occasional smokers also did not reveal an intervention effect on smoking abstinence. One explanation might be the short-term follow-up assessment, which was conducted 3 months after the end of the intervention. In motivational interventions addressing smokers irrespective of their intention to quit, the effects on smoking abstinence rates typically increase gradually [32,35] and might become statistically significant at later follow-up assessments.

The subgroup analyses revealed positive intervention effects for both subgroups on cigarette consumption. Furthermore, occasional smokers in the intervention group made more serious

attempts to quit smoking. Quit attempts are significant predictors of smoking cessation [32,36,37].

Several limitations must be noted. First, smoking status was assessed by self-report and was not biochemically verified. However, we expect that a potential overreporting of smoking abstinence would be independent of the study condition. Furthermore, based on recommendations by the Society for Research on Nicotine and Tobacco, there are circumstances under which the added precision gained by biological validation is offset in such a way that its use is not required and may not be desirable [38]. Examples include population-based studies with low demands on smokers to quit (eg, interventions with limited face-to-face contact and studies in which the optimal data collection methods are through mail, telephone, or Internet). A second limitation is that we only investigated the short-term effects of the program. Longer follow-up assessments might provide different results. However, both of these limitations resulted in a lower expenditure of time for the study participants and a greater proximity to prevention practice. Therefore, they allowed a better estimation of the participation rate in the program that might be expected under routine intervention conditions. Further study limitations are the lack of statistical

power, particularly for the subgroup analyses, and an attrition bias. Based on a higher percentage of daily smokers and higher cigarette consumption in individuals lost to follow-up as well as a higher percentage of persons lost to follow-up in the control group than in the intervention group, this attrition bias might have resulted in conservative estimations of intervention effects in the complete-case analyses.

The study demonstrates the potential of a text messaging-based intervention to reach a high proportion of young smokers with predominantly lower educational levels. The intervention resulted in statistically significant lower cigarette consumption in the total sample, the subgroup of occasional smokers, and the subgroup of daily smokers. Furthermore, it resulted in statistically significant more quit attempts in the subgroup of occasional smokers. No short-term effects were found according to the proportion of participants who had quit.

Both the baseline assessment and the registration for the SMS text message program are possible from every computer with Internet access and only take approximately 10 minutes. Therefore, the program could be easily implemented within school classes with low personnel expenses.

Acknowledgments

Funding for this study was provided by the Swiss Tobacco Prevention Fund (No: 10.007207). The funding institution had no role in the design or conduct of the study, the collection, management, analysis, or interpretation of the data, or the preparation, review, or approval of the manuscript. We would like to thank Dr Ahmet Turan Tagmat for the technical development of the program and the teachers and students of the vocational schools for supporting the project by their participation. We appreciate the work of the study assistants who collected the data.

Conflicts of Interest

Two authors (SH, CM) were involved in the development of the intervention.

Multimedia Appendix 1

Intervention components of the SMS program.

[[PPTX File, 1MB - jmir_v15i8e171_app1.pptx](#)]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [39].

[[PDF File \(Adobe PDF File\), 1000KB - jmir_v15i8e171_app2.pdf](#)]

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012 Dec 15;380(9859):2224-2260. [doi: [10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8)] [Medline: [23245609](https://pubmed.ncbi.nlm.nih.gov/23245609/)]
2. Hibell B, Guttormsson U, Ahlström S, Balakireva O, Bjarnason T, Kokkevi A, et al. The 2011 ESPAD report. Substance use among students in 36 European countries. Stockholm, Sweden: The Swedish Council for Information on Alcohol and other Drugs (CAN); 2012.
3. Hanke M, Ulbricht S, Freyer-Adam J, John U, Meyer C, Haug S. [Tobacco smoking and alcohol consumption among apprentices at vocational schools in West Pomerania, Germany]. *Gesundheitswesen* 2013 Apr;75(4):216-224. [doi: [10.1055/s-0032-1311621](https://doi.org/10.1055/s-0032-1311621)] [Medline: [22864848](https://pubmed.ncbi.nlm.nih.gov/22864848/)]
4. Grimshaw GM, Stanton A. Tobacco cessation interventions for young people. *Cochrane Database Syst Rev* 2006(4):CD003289. [doi: [10.1002/14651858.CD003289.pub4](https://doi.org/10.1002/14651858.CD003289.pub4)] [Medline: [17054164](https://pubmed.ncbi.nlm.nih.gov/17054164/)]

5. Sussman S, Sun P. Youth tobacco use cessation: 2008 update. *Tob Induc Dis* 2009;5:3 [FREE Full text] [doi: [10.1186/1617-9625-5-3](https://doi.org/10.1186/1617-9625-5-3)] [Medline: [19183452](https://pubmed.ncbi.nlm.nih.gov/19183452/)]
6. Hollis JF, Polen MR, Whitlock EP, Lichtenstein E, Mullooly JP, Velicer WF, et al. Teen reach: outcomes from a randomized, controlled trial of a tobacco reduction program for teens seen in primary medical care. *Pediatrics* 2005 Apr;115(4):981-989 [FREE Full text] [doi: [10.1542/peds.2004-0981](https://doi.org/10.1542/peds.2004-0981)] [Medline: [15805374](https://pubmed.ncbi.nlm.nih.gov/15805374/)]
7. Pbert L, Flint AJ, Fletcher KE, Young MH, Druker S, DiFranza JR. Effect of a pediatric practice-based smoking prevention and cessation intervention for adolescents: a randomized, controlled trial. *Pediatrics* 2008 Apr;121(4):e738-e747 [FREE Full text] [doi: [10.1542/peds.2007-1029](https://doi.org/10.1542/peds.2007-1029)] [Medline: [18381502](https://pubmed.ncbi.nlm.nih.gov/18381502/)]
8. Peterson AV, Kealey KA, Mann SL, Marek PM, Ludman EJ, Liu J, et al. Group-randomized trial of a proactive, personalized telephone counseling intervention for adolescent smoking cessation. *J Natl Cancer Inst* 2009 Oct 21;101(20):1378-1392 [FREE Full text] [doi: [10.1093/jnci/djp317](https://doi.org/10.1093/jnci/djp317)] [Medline: [19822836](https://pubmed.ncbi.nlm.nih.gov/19822836/)]
9. Glasgow RE, Klesges LM, Dziewaltowski DA, Estabrooks PA, Vogt TM. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Educ Res* 2006 Oct;21(5):688-694 [FREE Full text] [doi: [10.1093/her/cyl081](https://doi.org/10.1093/her/cyl081)] [Medline: [16945984](https://pubmed.ncbi.nlm.nih.gov/16945984/)]
10. Backinger CL, Michaels CN, Jefferson AM, Fagan P, Hurd AL, Grana R. Factors associated with recruitment and retention of youth into smoking cessation intervention studies--a review of the literature. *Health Educ Res* 2008 Apr;23(2):359-368 [FREE Full text] [doi: [10.1093/her/cym053](https://doi.org/10.1093/her/cym053)] [Medline: [17884834](https://pubmed.ncbi.nlm.nih.gov/17884834/)]
11. Bühler A, Wegmann L, Schmidt A, Thurl J, Strunk M, Lang P. Rekrutierung, Implementation und Aufhorräte eines Rauchstopp-Kurses für Jugendliche [Recruitment, implementation and quit rate of an adolescent smoking cessation program]. *SUCHT - Zeitschrift für Wissenschaft und Praxis [Journal of Addiction Research and Practice]* 2012 Feb 23;58(1):23-32. [doi: [10.1024/0939-5911.a000158](https://doi.org/10.1024/0939-5911.a000158)]
12. Willemsse I, Waller G, Süss D. JAMES: Jugend, Aktivität, Medien - Erhebung Schweiz [Youth, Activity, Media in Switzerland]. Zürich: Zürcher Hochschule für angewandte Wissenschaften; 2010. URL: http://www.psychologie.zhaw.ch/fileadmin/user_upload/psychologie/Downloads/Forschung/JAMES/Ergebnisbericht_JAMES_2010_de.pdf [accessed 2013-07-29] [WebCite Cache ID 6ITYobcrY]
13. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull* 2007 Jul;133(4):673-693. [doi: [10.1037/0033-2909.133.4.673](https://doi.org/10.1037/0033-2909.133.4.673)] [Medline: [17592961](https://pubmed.ncbi.nlm.nih.gov/17592961/)]
14. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2012;11:CD006611. [doi: [10.1002/14651858.CD006611.pub3](https://doi.org/10.1002/14651858.CD006611.pub3)] [Medline: [23152238](https://pubmed.ncbi.nlm.nih.gov/23152238/)]
15. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. *Lancet* 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: [10.1016/S0140-6736\(11\)60701-0](https://doi.org/10.1016/S0140-6736(11)60701-0)] [Medline: [21722952](https://pubmed.ncbi.nlm.nih.gov/21722952/)]
16. Haug S, Meyer C, Schorr G, Bauer S, John U. Continuous individual support of smoking cessation using text messaging: a pilot experimental study. *Nicotine Tob Res* 2009 Aug;11(8):915-923. [doi: [10.1093/ntr/ntp084](https://doi.org/10.1093/ntr/ntp084)] [Medline: [19542517](https://pubmed.ncbi.nlm.nih.gov/19542517/)]
17. Haug S, Meyer C, Gross B, Schorr G, Thyrian JR, Kordy H, et al. [Continuous individual support of smoking cessation in socially deprived young adults via mobile phones--results of a pilot study]. *Gesundheitswesen* 2008 Jun;70(6):364-371. [doi: [10.1055/s-2008-1080932](https://doi.org/10.1055/s-2008-1080932)] [Medline: [18661460](https://pubmed.ncbi.nlm.nih.gov/18661460/)]
18. Haug S, Bitter G, Hanke M, Ulbricht S, Meyer C, John U. [Brief Text Messaging (SMS)-Delivered Smoking Cessation Intervention for Apprentices in Vocational Schools: A Longitudinal Intervention Study for Process and Outcome Evaluation.]. *Gesundheitswesen* 2013 Jan 4. [doi: [10.1055/s-0032-1331728](https://doi.org/10.1055/s-0032-1331728)] [Medline: [23292965](https://pubmed.ncbi.nlm.nih.gov/23292965/)]
19. Haug S, Meyer C, Dymalski A, Lippke S, John U. Efficacy of a text messaging (SMS) based smoking cessation intervention for adolescents and young adults: study protocol of a cluster randomised controlled trial. *BMC Public Health* 2012;12:51 [FREE Full text] [doi: [10.1186/1471-2458-12-51](https://doi.org/10.1186/1471-2458-12-51)] [Medline: [22260736](https://pubmed.ncbi.nlm.nih.gov/22260736/)]
20. 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)]
21. Heatherton TF, Kozlowski LT, Frecker RC, Rickert W, Robinson J. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *Br J Addict* 1989 Jul;84(7):791-799. [Medline: [2758152](https://pubmed.ncbi.nlm.nih.gov/2758152/)]
22. Pocock SJ. *Clinical Trials: A Practical Approach*. Chichester: Wiley & Sons; 1994.
23. Perkins HW. *The Social Norms Approach to Preventing School and College Age Substance Abuse: A Handbook for Educators, Counselors, and Clinicians*. First ed. San Francisco, CA: Jossey-Bass; 2003.
24. 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: [10.1026/S-0065-2601\(06\)38002-1](https://doi.org/10.1026/S-0065-2601(06)38002-1)]
25. Lippke S, Ziegelmann JP, Schwarzer R, Velicer WF. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. *Health Psychol* 2009 Mar;28(2):183-193 [FREE Full text] [doi: [10.1037/a0012983](https://doi.org/10.1037/a0012983)] [Medline: [19290710](https://pubmed.ncbi.nlm.nih.gov/19290710/)]
26. Suppli CH, Due P, Henriksen PW, Rayce SL, Holstein BE, Rasmussen M. Low vigorous physical activity at ages 15, 19 and 27: childhood socio-economic position modifies the tracking pattern. *Eur J Public Health* 2013 Feb;23(1):19-24. [doi: [10.1093/eurpub/cks040](https://doi.org/10.1093/eurpub/cks040)] [Medline: [22552259](https://pubmed.ncbi.nlm.nih.gov/22552259/)]

27. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction* 1993 Jun;88(6):791-804. [Medline: [8329970](#)]
28. Bush K, Kivlahan DR, McDonnell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998 Sep 14;158(16):1789-1795. [Medline: [9738608](#)]
29. Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR. AUDIT-C as a brief screen for alcohol misuse in primary care. *Alcohol Clin Exp Res* 2007 Jul;31(7):1208-1217. [doi: [10.1111/j.1530-0277.2007.00403.x](#)] [Medline: [17451397](#)]
30. Morgenstern M, Herrmann U, Hanewinkel R. Rauchstopp-Hilfen für Jugendliche: Welche Faktoren beeinflussen Bekanntheit, Akzeptanz und Nutzung? Schlussbericht zum Forschungsprojekt [Smoking cessation aids for adolescents: Which factors affect awareness, acceptance and use? Research Report]. Kiel, Germany: Institut für Therapie- und Gesundheitsforschung; 2009.
31. Japuntich SJ, Piper ME, Schlam TR, Bolt DM, Baker TB. Do smokers know what we're talking about? The construct validity of nicotine dependence questionnaire measures. *Psychol Assess* 2009 Dec;21(4):595-607 [FREE Full text] [doi: [10.1037/a0017312](#)] [Medline: [19947792](#)]
32. Velicer WF, Redding CA, Sun X, Prochaska JO. Demographic variables, smoking variables, and outcome across five studies. *Health Psychol* 2007 May;26(3):278-287. [doi: [10.1037/0278-6133.26.3.278](#)] [Medline: [17500614](#)]
33. Haug S, Meyer C, Ulbricht S, Schorr G, Rüge J, Rumpf HJ, et al. Predictors and moderators of outcome in different brief interventions for smoking cessation in general medical practice. *Patient Educ Couns* 2010 Jan;78(1):57-64. [doi: [10.1016/j.pec.2009.07.005](#)] [Medline: [19660890](#)]
34. Kleinjan M, Engels RC, van Leeuwe J, Brug J, van Zundert RM, van den Eijnden RJ. Mechanisms of adolescent smoking cessation: roles of readiness to quit, nicotine dependence, and smoking of parents and peers. *Drug Alcohol Depend* 2009 Jan 1;99(1-3):204-214. [doi: [10.1016/j.drugalcdep.2008.08.002](#)] [Medline: [18848408](#)]
35. Meyer C, Ulbricht S, Baumeister SE, Schumann A, Rüge J, Bischof G, et al. Proactive interventions for smoking cessation in general medical practice: a quasi-randomized controlled trial to examine the efficacy of computer-tailored letters and physician-delivered brief advice. *Addiction* 2008 Feb;103(2):294-304 [FREE Full text] [doi: [10.1111/j.1360-0443.2007.02031.x](#)] [Medline: [17995993](#)]
36. Farkas AJ, Pierce JP, Zhu SH, Rosbrook B, Gilpin EA, Berry C, et al. Addiction versus stages of change models in predicting smoking cessation. *Addiction* 1996 Sep;91(9):1271-80; discussion 1281. [Medline: [8854358](#)]
37. Abrams DB, Herzog TA, Emmons KM, Linnan L. Stages of change versus addiction: a replication and extension. *Nicotine Tob Res* 2000 Aug;2(3):223-229. [Medline: [11082822](#)]
38. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. *Nicotine Tob Res* 2002 May;4(2):149-159. [doi: [10.1080/14622200210123581](#)] [Medline: [12028847](#)]
39. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](#)] [Medline: [22209829](#)]

Abbreviations

AUDIT-C: Alcohol Use Disorder Identification Test (consumption questions)

CCA: complete-case analyses

HAPA: Health Action Process Approach

HBSC: Health Behavior in School-Aged Children

ITT: intention-to-treat

OR: odds ratio

SMS: short message service

Edited by G Eysenbach; submitted 27.03.13; peer-reviewed by R Schwarzer, D Schulz; comments to author 21.04.13; revised version received 25.04.13; accepted 11.06.13; published 16.08.13.

Please cite as:

Haug S, Schaub MP, Venzin V, Meyer C, John U

Efficacy of a Text Message-Based Smoking Cessation Intervention for Young People: A Cluster Randomized Controlled Trial
J Med Internet Res 2013;15(8):e171

URL: <http://www.jmir.org/2013/8/e171/>

doi: [10.2196/jmir.2636](#)

PMID: [23956024](#)

©Severin Haug, Michael P Schaub, Vigeli Venzin, Christian Meyer, Ulrich John. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 16.08.2013. 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

Enhancing Retention of an Internet-Based Cohort Study of Men Who Have Sex With Men (MSM) via Text Messaging: Randomized Controlled Trial

Christine M Khosropour^{1,2}, MPH; Brent A Johnson³, PhD; Alexandra V Ricca², MPH; Patrick S Sullivan², DVM, PhD

¹School of Public Health, Department of Epidemiology, University of Washington, Seattle, WA, United States

²Rollins School of Public Health, Department of Epidemiology, Emory University, Atlanta, GA, United States

³Rollins School of Public Health, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, United States

Corresponding Author:

Patrick S Sullivan, DVM, PhD

Rollins School of Public Health

Department of Epidemiology

Emory University

Grace Crum Rollins Building, 4th Floor

1518 Clifton Road NE

Atlanta, GA, 30322

United States

Phone: 1 404 727 2038

Fax: 1 866 311 8234

Email: pssulli@emory.edu

Abstract

Background: Black and Hispanic men who have sex with men (MSM) are disproportionately affected by HIV in the United States. The Internet is a promising vehicle for delivery of HIV prevention interventions to these men, but retention of MSM of color in longitudinal Internet-based studies has been problematic. Text message follow-up may enhance retention in these studies.

Objective: To compare retention in a 12-month prospective Internet-based study of HIV-negative MSM randomized to receive bimonthly follow-up surveys either through an Internet browser online or through text messages.

Methods: Internet-using MSM were recruited through banner advertisements on social networking and Internet-dating sites. White, black, and Hispanic men who were ≥ 18 , completed an online baseline survey, and returned an at-home HIV test kit, which tested HIV negative, were eligible. Men were randomized to receive follow-up surveys every 2 months on the Internet or by text message for 12 months (unblinded). We used time-to-event methods to compare the rate of loss-to-follow-up (defined as non-response to a follow-up survey after multiple systematically-delivered contact attempts) in the 2 follow-up groups, overall and by race/ethnicity. Results are reported as hazard ratios (HR) and 95% confidence intervals (CI) of the rate of loss-to-follow-up for men randomized to text message follow-up compared to online follow-up.

Results: Of 1489 eligible and consenting men who started the online baseline survey, 895 (60%) completed the survey and were sent an at-home HIV test kit. Of these, 710 of the 895 (79%) returned the at-home HIV test kit, tested HIV-negative, and were followed prospectively. The study cohort comprised 66% white men (470/710), 15% (106/710) black men, and 19% (134/710) Hispanic men. At 12 months, 77% (282/366) of men randomized to online follow-up were retained in the study, compared to 70% (241/344) men randomized to text message follow-up (HR=1.30, 95% CI 0.97-1.73). The rate of loss-to-follow-up was non-significantly higher in the text message arm compared to the online arm for both white (HR=1.43, 95% CI 0.97-1.73) and Hispanic men (HR=1.71, 95% CI 0.91-3.23); however, loss-to-follow-up among black men was non-significantly lower among those who received text message follow-up compared to online follow-up (HR=0.78, 95% CI 0.41-1.50). In the online arm, black men were significantly more likely to be lost to follow-up compared to white men (HR=2.25, 95% CI 1.36-3.71), but this was not the case in the text message arm (HR=1.23, 95% CI 0.70-2.16).

Conclusions: We retained $>70\%$ of MSM enrolled in an online study for 12 months; thus, engaging men in online studies for a sufficient time to assess sustained outcomes is possible. Text message follow-up of an online cohort of MSM is feasible, and may result in higher retention among black MSM.

KEYWORDS

HIV infections/prevention and control; prospective studies; Internet/organization and administration; SMS text messaging; homosexuality; male/statistics and numerical data

Introduction

Over 60% of all new HIV diagnoses in the United States are among men who have sex with men (MSM) [1], who represent only an estimated 3%-7% of the US population [2]. From 2006-2009, HIV incidence increased by 34% among all young MSM aged 13-29, with an increased incidence of 48% noted among young black MSM [3]. Nearly three-quarters of the new HIV infections among young Hispanic Americans in 2009 were among MSM [3].

The recent increases in incidence among young MSM of color have led to a call for new approaches to HIV prevention [4], including technology-based prevention interventions [5]. The Internet is an attractive vehicle for intervention delivery for many reasons, including minimal cost relative to interventions utilizing human resources, standardization of intervention content [6], inclusion of high-risk MSM who may not be reached by in-person sampling methods [7,8], and recruitment of the large number of MSM required to use HIV incidence as a study endpoint [9,10]. Moreover, a recent meta-analysis showed that computer-delivered interventions are similarly efficacious to traditional, human-delivered interventions [6].

Despite the benefits of Internet-based interventions, retention in online cohort studies of MSM has been problematic. In three online studies of MSM, 3-month retention was between 15%-54% [11-13], below the 70% required by the Centers for Disease Control and Prevention (CDC) Prevention Research Synthesis criteria for best-evidence HIV prevention interventions [14]. Further, retention of black MSM in a number of online studies has been significantly lower than that of white MSM [12,13,15]; thus, results from these studies may not adequately represent those of black MSM or may accrue biases.

Differences in retention in online studies by race/ethnicity might be partially explained by differences in Internet access. In 2011, approximately 66% of white Americans had broadband Internet access in the household, compared to 49% of black and 51% of Hispanic Americans [16]. In contrast, mobile phone ownership among black and Hispanic Americans is equivalent to that of white Americans. National surveys conducted in 2012-2013 [17,18] indicate that 93% of black and 88% of Hispanic Americans owned a mobile phone compared to 90% of white Americans. Among mobile phone owners, a similar proportion of black and Hispanic Americans reported using text messaging (80% and 85%, respectively) compared to white Americans (79%), and 97% of young Americans aged 18-29 reported using their phones for SMS text messaging.

Because young black and Hispanic Americans are high users of mobile technology, we sought to investigate whether the use of text messaging would increase retention in a 12-month online cohort study of HIV-negative white, black, and Hispanic MSM. Our primary aims were to compare the 12-month retention of

MSM randomized to receive online follow-up surveys versus text message follow-up surveys and to compare 12-month retention by race/ethnicity. We hypothesized that providing follow-up surveys by text messaging would result in higher retention, especially among MSM of color. Additionally, we describe mobile access to online surveys and the frequency of changing mobile phone numbers.

Methods

Study Design and Population

MSM were recruited from August to December 2010 by banner advertisements placed on social networking and select Internet-dating websites, including Facebook, Myspace, Black Gay Chat, and Adam4Adam. Website selection was based on data from four focus groups of MSM conducted in 2010. We chose websites that men indicated that they visited frequently or sites on which they felt the advertisements were "legitimate" and "trustworthy". Eligible participants were male, at least 18 years of age, white non-Hispanic, black non-Hispanic, or Hispanic, and reported sex with a man in the past 12 months. Additional eligibility criteria included owning a mobile phone capable of sending and receiving text messages, being willing to receive an at-home HIV test kit, and not moving outside the United States in the next 12 months. Because we were interested in determining the retention of an Internet-based sample of HIV-negative MSM, only those men who returned their HIV test kit and tested HIV-negative were followed prospectively for 12 months.

Men provided electronic informed consent prior to initiating any study procedures by checking a box on the survey screen. Consenting men were asked to register for the study by providing an email address and mobile phone number before completing the baseline survey. These were validated in two sequential steps. First, a unique participant-specific URL for the baseline survey was sent to the participant's email address. Second, participants who successfully linked into the baseline survey through the URL in their email were asked to enter their mobile phone number, to which a 3-digit code was sent by text message. Participants entered the 3-digit code on a survey screen in order to proceed in the study.

Men with verified email and mobile phone information completed a 60-minute baseline survey that included questions on condom acquisition and use, demographics, sexual risk behaviors, sexual partner history, and HIV testing history. At the conclusion of the baseline survey, men who did not report being HIV-positive provided their mailing address for an at-home HIV test kit. Those who provided a valid mailing address were randomized 1:1 to receive either text message or online follow-up surveys every 2 months for a total of 12 months. Randomization was implemented through the online enrollment system; there were no blocks of randomization, so

men were assigned to an arm through random number generation at the time each man was determined to be eligible. Participants were not blinded to the arm to which they were randomized. To facilitate completion of the follow-up surveys, we asked participants to choose a preferred day of the week and time of day to receive their follow-up surveys. Additionally, we requested that participants indicate a preferred alternate contact method, in the event that we were unable to contact them via email (for the online arm) or text message (for the text message arm).

Participants were compensated US \$15 for completing the baseline survey, US \$10 for each follow-up survey, and US \$15 for the Month 12 survey. Men were also paid US \$20 for returning their at-home test kit. Payments were delivered via PayPal or Amazon.com electronic gift card after completion of each survey. Participants randomized to the text message arm who did not have a text message plan from their mobile phone carrier were charged US \$0.10 per text message response. The cost for text messages sent to participants was paid by the research team. Before providing informed consent, potential participants were informed about the potential to incur costs associated with sending text messages as part of the study.

Follow-Up

Participants received notifications to take their follow-up surveys 8 weeks after their last completed survey. Participants randomized to receive online follow-up surveys received an email that contained a unique URL to link to the follow-up survey. Participants randomized to receive text message follow-up surveys received a text message that provided an opportunity for participants to initiate the survey immediately or delay the survey for 24 hours. The text message survey was a question-and-response format (ie, the subsequent survey question was only sent once the response to the previous question had been received). Similar to the online survey, the text message survey incorporated skip patterns based on participant responses so that only relevant questions were asked. The content of the follow-up surveys, which queried men on their 2-month sexual history and HIV testing history, was identical for both randomization arms. Regardless of randomization arm, all participants received an email notification for the final (Month 12) survey, which was administered online.

We used a systematic arm-dependent method to maximize retention. Men randomized to the online arm who had not completed the survey 3 days after the initial notification email were automatically sent a reminder email. Two subsequent automated reminder emails were then sent, each separated by 24 hours. Men randomized to the text message arm who did not initially complete the survey or did not request a delay of survey initiation received 3 additional text message reminders, each separated by 24 hours. Men in both randomization arms who did not complete a follow-up survey after the first group of reminders were contacted up to 3 additional times by study staff, using the preferred method of contact provided in the baseline survey. As a final step, study staff called the participant via mobile phone to remind him to complete his follow-up survey.

Participants were withdrawn from the study if they did not complete the follow-up survey after 3 phone calls.

Outcome

The primary outcome was loss to follow-up, defined as administrative withdrawal by study staff (for non-response, as described above), before the Month 12 survey, or request by a participant to be withdrawn from the study.

Statistical Analysis

Using methods for time-to-event data [19,20], we defined the period of analysis as the date of randomization until (1) the earliest of 365 days post-randomization or the date of completion of the Month 12 survey (for participants who were retained in the study); or (2) the date of the most recently completed survey (for participants who were lost to follow-up). Consequently, participants who were retained in the study but had not completed the final survey at the end of the analysis period (ie, 365 days after randomization) were considered censored.

Descriptive statistics were used to assess the distribution of participant characteristics by randomization arm, stratified by race/ethnicity. We used the Kaplan-Meier estimator to examine the rate of loss-to-follow-up by randomization arm and by race/ethnicity. We used Cox proportional hazards regression to estimate the hazard ratio (HR) and 95% confidence interval (CI) of time to loss-to-follow-up associated with randomization arm, overall, and stratified by race/ethnicity. We also estimated the HR and corresponding 95% CI of the rate of loss-to-follow-up within randomization arm for black and Hispanic participants relative to white participants.

We used scaled Schoenfeld residuals to evaluate the proportional hazards assumption of the Cox regression models [21]. For the primary model comparing randomization arms, a formal statistical test rejected the hypothesis of proportional hazards ($P<.001$). We determined that the relative hazard ratio changed sign at about 300 days (see [Multimedia Appendix 1](#)), which is consistent with the final time that men who were randomized to text message follow-up completed a text message survey. Therefore, we report the Cox regression estimates comparing online and text message follow-up for two models: one based on all data through Month 12 (365 days) and one based on data up to and including 300 days (Month 10). We did not detect a departure from the proportional hazards assumption for the model comparing retention by racial/ethnic group ($P=.59$); thus, we only report results using all data through Month 12 for that analysis. Reported P -values for all analyses are based on the Wald test of significance (alpha=0.05 level). Analyses were conducted in R and Stata 12.1.

All study procedures and analysis were reviewed and approved by the Institutional Review Board (IRB) of Emory University. This study used a randomized method for follow-up but did not meet the qualifications for ClinicalTrials.gov registration (ie, the study did not “prospectively assign human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes”.) This study did not utilize a health-related intervention or a health outcome.

Results

Study Population

There were 6174 MSM eligible to participate in the study (Figure 1). Less than half of all consenting MSM (1489/3474) provided valid contact information and initiated the baseline survey. Of those who completed the baseline survey and were randomized (n=895), return of the at-home test kit was similar by randomization arm (81% online vs 78% text message, $P=.34$). In total, 710 MSM tested HIV-negative and were sent bimonthly follow-up surveys for 12 months.

Of 710 participants, 366 (52%) were randomized to online follow-up and 344 (49%) were randomized to text message follow-up. Two-thirds of participants were white and slightly over one-third were ≤ 24 years old (Table 1). Most men had at least some college-level education, and two-thirds resided in an urban area. Characteristics of participants between randomization arm did not differ within racial/ethnic groups. Characteristics were also balanced among men initially randomized to follow-up (n=895).

Retention by Follow-Up Arm and Racial/Ethnic Group

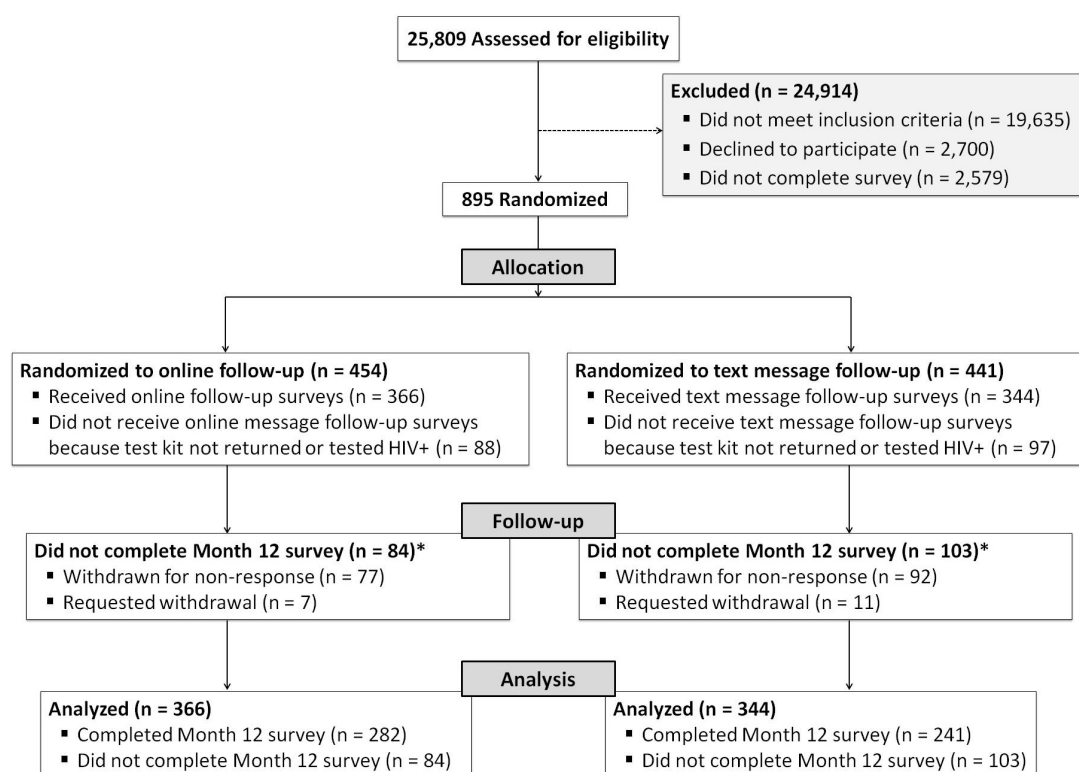
Overall, 74% of men (523/710) were retained in the study at 12 months. Of the 187 men who were lost to follow-up, 18 (10%) requested to be withdrawn from the study and 169 (90%) were administratively withdrawn due to non-response. Withdrawal requests were more common among white men than black and Hispanic men: 13% (14/108) of white men, 8% (3/38) of black men, and 2% (1/41) of Hispanic men ($P=.14$) who were lost to follow-up requested to be withdrawn from the study (data not shown).

Nearly 10% (65/710) of men were lost to follow-up before the first follow-up survey (Figure 2). At 12 months, men randomized to text message follow-up had a higher rate of loss-to-follow-up compared to men randomized to online follow-up, although this difference was not statistically significant (Table 2). Among white and Hispanic men, being randomized to text message follow-up was associated with a 42% and 71% higher rate of loss-to-follow-up, respectively, compared to online follow-up. In contrast, black men randomized to text message follow-up had a 20% reduction in the rate of loss-to-follow-up compared to black men randomized to online follow-up. Results using data only through Month 10 (300 days) did not differ (Table 2).

Black men were less likely to be retained in the study compared to white or Hispanic men (Figure 3). Among men randomized to online follow-up, black men had a two-fold higher rate of loss-to-follow-up compared to white men (Table 3). However, there was no significant difference in the rate of loss-to-follow-up between black and white men randomized to text message follow-up. Compared to white men, Hispanic men randomized to text message follow-up had a somewhat higher rate of loss-to-follow-up.

Approximately 20% (71/362) of white men, 22% (15/68) of black men, and 17% (16/93) of Hispanic men who completed the Month 12 online survey did so on a mobile browser. Additionally, in a 5-month period beginning in February 2011, 27 of 244 (8%) participants in the text message arm notified study staff that they had acquired new mobile phone numbers. This included 6% (14/229) of white participants, 11% (5/46) of black participants, and 12% (8/69) of Hispanic participants.

Figure 1. Enrollment of study participants in a 12-month prospective online study.



*Loss to follow-up is the primary outcome of interest in the study; thus, these men are included in the "Analysis" boxes that follow

Table 1. Baseline characteristics of participants in an online study, by race/ethnicity and randomization arm (N=710).^{a, b}

Characteristic	Total (N=710), n (%)	White		Black		Hispanic		
		Online (N=241), n (%)	SMS (N=229), n (%)	Online (N=60), n (%)	SMS (N=46), n (%)	Online (N=65), n (%)	SMS (N=69), n (%)	
Age group								
18-24	263 (37.0)	71 (29.5)	79 (34.5)	28 (46.7)	25 (54.4)	28 (43.1)	32 (46.4)	
25-34	262 (36.9)	94 (39.0)	81 (35.4)	20 (33.3)	16 (34.8)	27 (41.5)	24 (34.8)	
35-44	107 (15.1)	43 (17.8)	39 (17.0)	7 (11.7)	4 (8.7)	6 (9.2)	8 (11.6)	
45-54	78 (11.0)	33 (13.7)	30 (13.1)	5 (8.3)	1 (2.2)	4 (6.2)	5 (7.3)	
Education								
≤ High school	130 (18.3)	41 (17.0)	42 (18.3)	14 (23.3)	5 (10.9)	10 (15.4)	18 (26.1)	
> High school / GED	580 (82.7)	200 (83.0)	187 (81.7)	46 (76.7)	41 (89.1)	55 (84.6)	51 (73.9)	
Geographic region^c								
West	197 (27.7)	56 (23.2)	63 (27.5)	9 (15.0)	4 (8.7)	25 (38.5)	40 (58.0)	
Midwest	96 (13.5)	41 (17.0)	32 (14.0)	4 (6.7)	6 (13.0)	6 (9.2)	7 (10.1)	
South	278 (39.1)	87 (36.0)	90 (39.3)	40 (66.7)	26 (56.5)	19 (29.2)	16 (23.2)	
Northeast	139 (19.6)	57 (23.7)	44 (19.2)	7 (11.7)	10 (21.7)	15 (23.1)	6 (8.7)	
Residence^d								
Urban ^e	459 (66.6)	150 (63.0)	141 (63.2)	47 (81.0)	31 (73.8)	46 (74.2)	44 (66.7)	
Rural	230 (33.4)	88 (37.0)	82 (36.8)	11 (19.0)	11 (26.2)	16 (25.8)	22 (33.3)	
Sexual identity								
Homosexual	603 (84.9)	218 (90.5)	201 (87.8)	40 (66.7)	31 (67.4)	53 (81.5)	60 (87.0)	
Bisexual	88 (12.4)	20 (8.3)	23 (10.0)	16 (26.7)	9 (19.6)	12 (18.5)	8 (11.6)	
Other	19 (2.7)	3 (1.2)	5 (2.2)	4 (6.7)	6 (13.0)	0 (0.0)	1 (1.5)	
Ever tested for HIV	556 (78.5)	195 (81.3)	178 (78.1)	45 (75.0)	39 (84.8)	52 (80.0)	47 (68.1)	
HIV test in past 12m	337 (47.6)	118 (49.2)	104 (45.6)	26 (43.3)	24 (52.2)	35 (53.9)	30 (43.5)	
Sex of SP, past 12m								
Men	660 (93.0)	229 (95.0)	216 (94.3)	55 (91.7)	36 (78.3)	60 (92.3)	64 (92.8)	
Men and women	50 (7.0)	12 (5.0)	13 (5.7)	5 (8.3)	10 (21.7)	5 (7.7)	5 (7.3)	
No. of MSP past 12m								
1	102 (14.5)	34 (14.1)	33 (14.4)	5 (8.6)	7 (15.6)	9 (14.1)	14 (20.3)	
2-5	300 (42.5)	99 (41.1)	103 (45.0)	27 (46.6)	19 (42.2)	25 (39.1)	27 (39.1)	
>5	304 (43.1)	108 (44.8)	93 (40.6)	26 (44.8)	19 (42.2)	30 (46.9)	28 (40.6)	
UAI with MSP, past 12m	556 (83.0)	196 (86.0)	181 (82.7)	42 (73.7)	33 (82.5)	50 (80.7)	54 (84.4)	

^aOwing to missing data, numbers may not sum to column total. Denominators for proportions include those without missing data for that characteristic.

^bAbbreviations—GED: general equivalency diploma; 12m: 12 months; (M)SP: (male) sex partner; UAI: unprotected anal intercourse.

^cAs defined by the US Census Bureau.

^dBased on zip code where participant requested that at-home HIV test kit was sent.

^eUrban defined as residence in a zip code with population ≥1000 per square mile.

Figure 2. Retention of participants in an online study, by randomization arm (N=710).

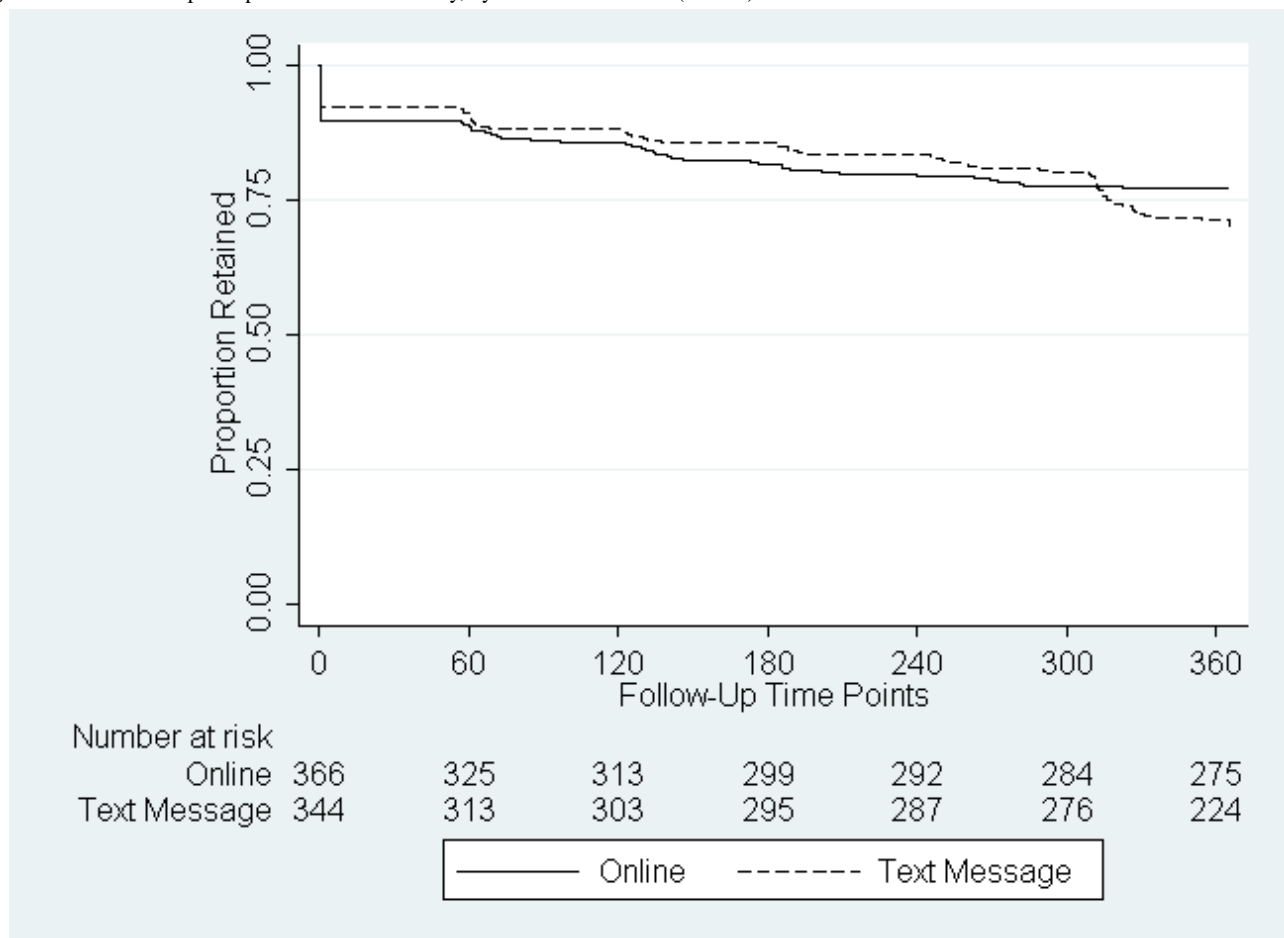


Table 2. Rate of loss-to-follow-up among men participating in a 12-month online study randomized to text message versus online follow-up, overall and stratified by race/ethnicity (N=710).

		Proportion retained at 12 months n (%) ^a	Month 12 estimates		Month 10 estimates	
			Hazard ratio	95% CI	Hazard ratio	95% CI
			Overall	Online	282/366 (77.1)	Referent
	Text message	241/344 (70.1)	1.30	0.97-1.73	1.28	0.96-1.71
White	Online	195/241 (80.9)	Referent	—	Referent	—
	Text message	167/229 (72.9)	1.43	0.97-2.09	1.41	0.96-2.06
Black	Online	37/60 (61.7)	Referent	—	Referent	—
	Text message	31/46 (67.4)	0.78	0.41-1.50	0.76	0.40-1.46
Hispanic	Online	50/65 (76.9)	Referent	—	Referent	—
	Text message	43/69 (62.3)	1.71	0.91-3.23	1.67	0.89-3.16

^aNumber retained out of total number defined by row.

Figure 3. Retention of participants in an online study, by race/ethnicity (N=710).

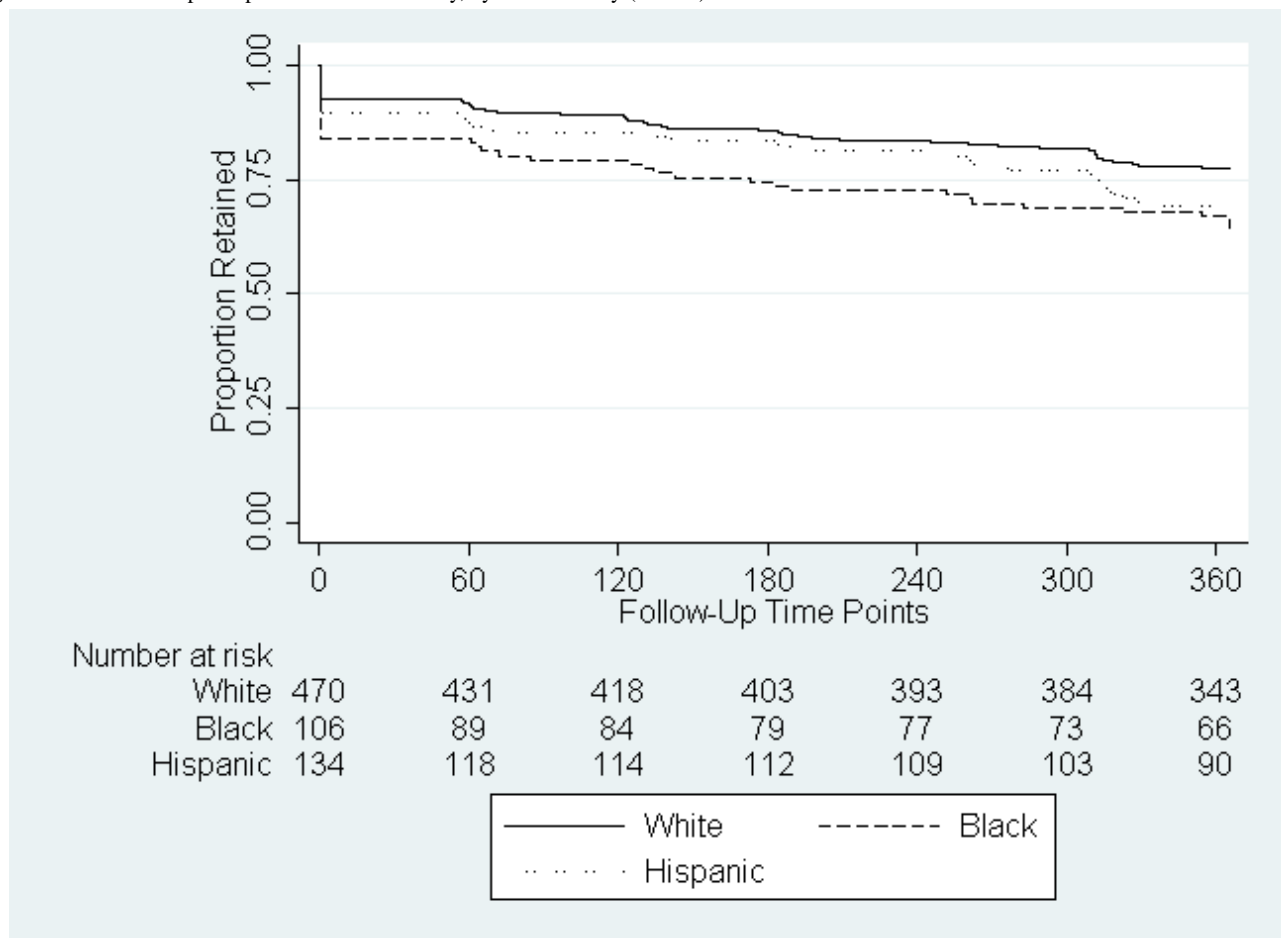


Table 3. Rate of loss-to-follow-up among men participating in a 12-month online study randomized to text message versus online follow-up, by race/ethnicity and stratified by randomization arm (N=710).

		Proportion retained at 12 months, n (%) ^a	Hazard ratio	95% CI
Overall				
	White	362/470 (77.0)	Referent	—
	Black	68/106 (64.0)	1.70	1.17-2.45
	Hispanic	93/134 (69.4)	1.37	0.96-1.97
Online follow-up				
	White	195/241 (80.9)	Referent	—
	Black	37/60 (61.7)	2.25	1.36-3.71
	Hispanic	50/65 (76.9)	1.22	0.68-2.19
Text message follow-up				
	White	167/229 (72.9)	Referent	—
	Black	31/46 (67.4)	1.23	0.70-2.16
	Hispanic	43/69 (62.3)	1.47	0.94-2.35

^aNumber retained out of total number defined by row

Discussion

Principal Findings

In this large Internet-based cohort of MSM, nearly three-quarters of men were retained in the study for 12 months. Black men

randomized to text message follow-up were somewhat more likely to be retained than those randomized to online follow-up, but this was not the case for white or Hispanic men. We observed a similar rate of loss-to-follow-up among black and white men randomized to text message follow-up, but black men followed exclusively online were less likely to be retained

than white men. Among all men, about 1 in 5 completed an online survey via mobile phone browser.

Our demonstrated retention of 74% at 12 months is encouraging to investigators wishing to establish large Internet-based cohorts of MSM. This proportion retained is over three times that observed by our group in a 3-month prospective of MSM in 2009 [12] and greater than that reported from other online studies of MSM using 2-month [15], 3-month [11,13], or 6-month [22] study endpoints. Our retention is somewhat lower than that observed in a 2007 Internet-based study by Horvath et al [23], in which 85% of MSM were retained. Although our study and that of Horvath et al were similar in their approach and design, differences in the recruitment strategy between that study (ie, website banner advertisements and email contact to men who had previously participated in a study) and our study (ie, website banner advertisement only) may explain the lower retention among our study participants. Notably, in our study and the study of Horvath et al, loss-to-follow-up occurred most often in the period immediately following baseline and stabilized somewhat thereafter.

The high retention in our study can be attributed to a number of factors. First, we validated contact information from study participants, thereby excluding men who would have been lost to follow-up due to erroneous email addresses or mobile phone numbers. Second, based on results from our previous study which identified factors associated with retention in an online cohort of MSM [12], we encouraged participants to provide an email address that they checked daily. Third, we allowed participants to choose the day of week and time of day that they would like to be contacted to complete the follow-up surveys, which may have increased the convenience of survey completion. Fourth, we used a systematic follow-up method to encourage participants to complete their follow-up surveys. As part of our reminder protocol, we collected alternate contact information for all study participants so that changes in email address or mobile phone numbers would not result in a loss of contact with the participant. Finally, our study cohort included men who had completed the baseline survey and returned an at-home HIV test kit. Therefore, we selected for study participants who were actively engaged in the research study and more likely to be retained for the duration of the study.

Given the demonstrated need for new HIV prevention interventions among black MSM, it is promising that nearly two-thirds of black study participants were retained in this study. Although our 12-month retention of black men in this study is lower than the 3-month retention (78%) observed by Hightow-Weidmen et al at 3 months [24], the retention we observed is considerably higher than the comparable proportion in other prospective online studies of black MSM recruited exclusively online [12,13] and is close to the 70% set by the CDC's criteria for best-evidence HIV prevention interventions [14]. The fact that black men randomized to text message follow-up had a higher retention than those followed exclusively online argues for the use of mobile measurement technologies to enhance research engagement in this group. Indeed, we noted a similar retention among black and white men in the text message arm, despite the relatively high loss-to-follow-up among black versus white men in the online follow-up. This

latter observation is consistent with previous studies of MSM followed exclusively online [9,12,13,15] and highlights the challenges in equalizing retention by race/ethnicity in Internet-based settings where participants are not provided mobile-based options for data collection.

We were surprised that Hispanic men randomized to text message follow-up did not demonstrate a higher retention than those randomized to online follow-up, considering the high mobile phone usage by Hispanic Americans and relatively low household Internet access [16,18]. We speculate that the low retention in the text message arm may be partially explained by switching mobile phone numbers; 12% of Hispanic study participants contacted study staff within a 5-month period to inform us of a mobile phone number change, but we received no such notifications for email address changes. However, the proportion of changed mobile phone numbers was similar for Hispanic and black men; therefore, this does not completely explain the low retention we observed among Hispanic men who received text messages.

We noted that one-fifth of participants completed the Month 12 survey on their mobile phone. If the proportion of bimonthly online follow-up surveys completed via mobile Web is similar to that of the Month 12 survey, the higher retention in the online arm may be partially explained by mobile access to the survey. There are two important implications of this finding. First, research studies wishing to use Internet-based data collection may benefit from employing mobile-enabled Internet surveys. Second, studies should consider offering multiple methods of data collection. It is possible that men who completed surveys via mobile Web did so because they did not have household Internet access or because they preferred the convenience of a mobile phone. Either way, retention in studies may be enhanced by allowing participants to choose their preferred method of technology.

Strengths and Limitations

This study has a number of strengths. We enrolled a large geographically and ethnically diverse cohort of MSM recruited from multiple websites. We employed a novel text message data collection system that incorporated skip patterns and recognition of invalid data entries. We used an automated reminder survey system that delivered surveys at specific times requested by study participants and sent automatic reminders at set time intervals.

There are several limitations to this analysis. First, we defined our study sample based on an event (return of the at-home HIV test kit) that occurred after randomization. Therefore, we may have lost some of the benefit of randomization to balance arms on confounding factors. Although characteristics of our study population were relatively similar by arm within racial/ethnic group, we cannot assess the distribution of unmeasured confounders. Second, our final study population included men who completed the baseline survey, provided valid contact information, and returned an at-home HIV test kit. Therefore, our population likely represents an actively engaged sample of research participants for which retention may be optimized. Third, although we specifically targeted websites to enhance recruitment of minority MSM (eg, Black Gay Chat), enrollment

of black and Hispanic men was below that of white men. This was disappointing, given that the goal of this study was to assess retention in an Internet-based cohort of minority men. However, this was not unanticipated, as we have previously characterized the under-enrollment of black and Hispanic MSM in online research [9]. Probability-based sampling has the potential to address the second and third limitations, but validly implementing a rigorous, probability-based sampling scheme over the Internet is challenging. Fourth, data on usage of a mobile phone browser were systematically collected only for the Month 12 survey. Therefore, the extent to which men accessed the online survey on their mobile phone for the bimonthly surveys is unknown. Fifth, men in this study are not representative of MSM who do not use social networking or Internet dating sites, or who do not click on advertisements displayed on these sites. Finally, our auxiliary statistical analysis of retention rates ([Multimedia Appendix 1](#)) suggested that a

time-varying coefficient Cox model (ie, one that allows the relative hazard ratio to fluctuate over time) may be more appropriate in future online studies. We addressed this potential limitation in the current analysis by analyzing and presenting all data as well as the subset of data that satisfied the proportional hazards assumption.

Conclusions

In summary, we demonstrated an ability to retain >70% of MSM enrolled in an online study for 12 months. Our study suggests that follow-up via text message is feasible and may result in higher retention among black MSM. Based on our findings, it is possible to engage MSM at greatest risk for HIV infection in large prospective, Internet-based HIV prevention intervention studies using a time interval that is sufficient to assess sustained outcomes [25].

Acknowledgments

The authors would like to thank the men who participated in this study. We also thank Nicole Luisi for her data management support and Leonard Linde for his development and management of the text message system. The author(s) disclose receipt of the following financial support for the research and/or authorship of this article: National Center for Minority Health and Health Disparities (1RC1MD004370) and Emory Center for AIDS Research (P30 AI050409).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Coefficient estimate for randomization arm effect: online versus text message.

[\[PDF File \(Adobe PDF File\), 95KB - jmir_v15i8e194_app1.pdf \]](#)

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [25].

[\[PDF File \(Adobe PDF File\), 993KB - jmir_v15i8e194_app2.pdf \]](#)

References

1. Centers for Disease Control and Prevention. HIV Surveillance Report, 2011; vol 23. 2013. URL: http://www.cdc.gov/hiv/pdf/statistics_2011_HIV_Surveillance_Report_vol_23.pdf [accessed 2013-05-01] [[WebCite Cache ID 6GIC0zS6F](#)]
2. Purcell DW, Johnson CH, Lansky A, Prejean J, Stein R, Denning P, et al. Estimating the population size of men who have sex with men in the United States to obtain HIV and syphilis rates. *Open AIDS J* 2012;6:98-107 [[FREE Full text](#)] [doi: [10.2174/1874613601206010098](https://doi.org/10.2174/1874613601206010098)] [Medline: [23049658](#)]
3. Prejean J, Song R, Hernandez A, Ziebell R, Green T, Walker F, HIV Incidence Surveillance Group. Estimated HIV incidence in the United States, 2006-2009. *PLoS One* 2011;6(8):e17502 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0017502](https://doi.org/10.1371/journal.pone.0017502)] [Medline: [21826193](#)]
4. Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. *Lancet* 2012 Jul 28;380(9839):388-399. [doi: [10.1016/S0140-6736\(12\)60955-6](https://doi.org/10.1016/S0140-6736(12)60955-6)] [Medline: [22819659](#)]
5. Sullivan PS, Grey JA, Simon Rosser BR. Emerging technologies for HIV prevention for MSM: what we have learned, and ways forward. *J Acquir Immune Defic Syndr* 2013 Jun 1;63 Suppl 1:S102-S107. [doi: [10.1097/QAI.0b013e3182949e85](https://doi.org/10.1097/QAI.0b013e3182949e85)] [Medline: [23673879](#)]
6. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS* 2009 Jan 2;23(1):107-115. [doi: [10.1097/QAD.0b013e32831c5500](https://doi.org/10.1097/QAD.0b013e32831c5500)] [Medline: [19050392](#)]

7. Sanchez T, Smith A, Denson D, Dinunno E, Lansky A. Internet-based methods may reach higher-risk men who have sex with men not reached through venue-based sampling. *Open AIDS J* 2012;6:83-89 [FREE Full text] [doi: [10.2174/1874613601206010083](https://doi.org/10.2174/1874613601206010083)] [Medline: [23049657](https://pubmed.ncbi.nlm.nih.gov/23049657/)]
8. Elford J, Bolding G, Davis M, Sherr L, Hart G. Web-based behavioral surveillance among men who have sex with men: a comparison of online and offline samples in London, UK. *J Acquir Immune Defic Syndr* 2004 Apr 1;35(4):421-426. [Medline: [15097159](https://pubmed.ncbi.nlm.nih.gov/15097159/)]
9. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. *J Med Internet Res* 2011;13(2):e38 [FREE Full text] [doi: [10.2196/jmir.1797](https://doi.org/10.2196/jmir.1797)] [Medline: [21571632](https://pubmed.ncbi.nlm.nih.gov/21571632/)]
10. Calabrese SK, Rosenberger JG, Schick VR, Novak DS, Reece M. An event-level comparison of risk-related sexual practices between black and other-race men who have sex with men: condoms, semen, lubricant, and rectal douching. *AIDS Patient Care STDS* 2013 Feb;27(2):77-84. [doi: [10.1089/apc.2012.0355](https://doi.org/10.1089/apc.2012.0355)] [Medline: [23373663](https://pubmed.ncbi.nlm.nih.gov/23373663/)]
11. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. *AIDS Care* 2004 Nov;16(8):931-943. [doi: [10.1080/09540120412331292507](https://doi.org/10.1080/09540120412331292507)] [Medline: [15511725](https://pubmed.ncbi.nlm.nih.gov/15511725/)]
12. Khosropour CM, Sullivan PS. Predictors of retention in an online follow-up study of men who have sex with men. *J Med Internet Res* 2011;13(3):e47 [FREE Full text] [doi: [10.2196/jmir.1717](https://doi.org/10.2196/jmir.1717)] [Medline: [21745792](https://pubmed.ncbi.nlm.nih.gov/21745792/)]
13. Chiasson MA, Shaw FS, Humberstone M, Hirshfield S, Hartel D. Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). *AIDS Care* 2009 Sep;21(9):1081-1089. [doi: [10.1080/09540120902730013](https://doi.org/10.1080/09540120902730013)] [Medline: [20024766](https://pubmed.ncbi.nlm.nih.gov/20024766/)]
14. Centers for Disease Control and Prevention. PRS Efficacy Criteria for Best-Evidence (Tier I) Risk Reduction (RR) Individual-level and Group-level Interventions (ILIs/GLIs). 2011. URL: http://www.cdc.gov/hiv/dhap/prb/prs/efficacy/rr/criteria/bestEvidence_ILIs_GLIs.html [accessed 2013-05-01] [WebCite Cache ID 6GIAIVhAg]
15. Bull SS, Vallejos D, Levine D, Ortiz C. Improving recruitment and retention for an online randomized controlled trial: experience from the Youthnet study. *AIDS Care* 2008 Sep;20(8):887-893. [doi: [10.1080/09540120701771697](https://doi.org/10.1080/09540120701771697)] [Medline: [18777217](https://pubmed.ncbi.nlm.nih.gov/18777217/)]
16. Zickuhr R, Smith A. Digital Differences. 2012. URL: http://www.pewinternet.org/~media/Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed 2013-07-04] [WebCite Cache ID 6HrjoBdC0]
17. Duggan M, Rainie L. Cell Phone Activities. 2012. URL: http://www.pewinternet.org/~media/Files/Reports/2012/PIP_CellActivities_11.25.pdf [accessed 2013-07-04] [WebCite Cache ID 6HrjgXo5o]
18. Brenner J. Pew Internet: Mobile. 2013. URL: <http://www.pewinternet.org/Commentary/2012/February/Pew-Internet-Mobile.aspx> [accessed 2013-07-04] [WebCite Cache ID 6Hrjji1T7]
19. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 1958;53(282):457-481.
20. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society* 1972;34(2):187-220.
21. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 1994;81(3):515-526.
22. Davidovich U, de Wit JBF, Stroebe W. Using the Internet to reduce risk of HIV-infection in steady relationships: a randomized controlled trial of a tailored intervention for gay men. In: Davidovich E, editor. *Liaisons dangereuses: HIV risk behavior prevention in steady gay relationships*. Amsterdam: Roel & Uigeefprojecten; 2006:95-122.
23. Horvath KJ, Nygaard K, Danilenko GP, Goknur S, Oakes JM, Rosser BR. Strategies to retain participants in a long-term HIV prevention randomized controlled trial: lessons from the MINTS-II study. *AIDS Behav* 2012 Feb;16(2):469-479 [FREE Full text] [doi: [10.1007/s10461-011-9957-3](https://doi.org/10.1007/s10461-011-9957-3)] [Medline: [21538084](https://pubmed.ncbi.nlm.nih.gov/21538084/)]
24. Hightow-Weidman LB, Pike E, Fowler B, Matthews DM, Kibe J, McCoy R, et al. HealthMpowerment.org: feasibility and acceptability of delivering an internet intervention to young Black men who have sex with men. *AIDS Care* 2012;24(7):910-920. [doi: [10.1080/09540121.2011.647677](https://doi.org/10.1080/09540121.2011.647677)] [Medline: [22272759](https://pubmed.ncbi.nlm.nih.gov/22272759/)]
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/)]

Abbreviations

CDC: Centers for Disease Control and Prevention
IRB: Institutional Review Board
MSM: men who have sex with men

Edited by G Eysenbach; submitted 07.06.13; peer-reviewed by K MacDonell, E Vandewater, L Hightow-Weidman; comments to author 02.07.13; revised version received 10.07.13; accepted 11.07.13; published 27.08.13.

Please cite as:

Khosropour CM, Johnson BA, Ricca AV, Sullivan PS

Enhancing Retention of an Internet-Based Cohort Study of Men Who Have Sex With Men (MSM) via Text Messaging: Randomized Controlled Trial

J Med Internet Res 2013;15(8):e194

URL: <http://www.jmir.org/2013/8/e194/>

doi: [10.2196/jmir.2756](https://doi.org/10.2196/jmir.2756)

PMID: [23981905](https://pubmed.ncbi.nlm.nih.gov/23981905/)

©Christine M Khosropour, Brent A Johnson, Alexandra V Ricca, Patrick S Sullivan. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.08.2013. 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 Twitter Among Local Health Departments: An Analysis of Information Sharing, Engagement, and Action

Brad L Neiger¹, PhD; Rosemary Thackeray¹, PhD; Scott H Burton², PhD; Callie R Thackeray¹, BS(Current); Jennifer H Reese¹, BS

¹Brigham Young University, Department of Health Science, Provo, UT, United States

²Brigham Young University, Department of Computer Science, Provo, UT, United States

Corresponding Author:

Brad L Neiger, PhD

Brigham Young University

Department of Health Science

302 B Widtsoe Building

Provo, UT, 84602

United States

Phone: 1 801 422 3313

Fax: 1 801 422 0050

Email: neiger@byu.edu

Abstract

Background: Social media offers unprecedented opportunities for public health to engage audiences in conversations and collaboration that could potentially lead to improved health conditions. While there is some evidence that local health departments (LHDs) are using social media and Twitter in particular, little is known about how Twitter is used by LHDs and how they use it to engage followers versus disseminating one-way information.

Objective: To examine how LHDs use Twitter to share information, engage with followers, and promote action, as well as to discover differences in Twitter use among LHDs by size of population served.

Methods: The Twitter accounts for 210 LHDs were stratified into three groups based on size of population served (n=69 for less than 100,000; n=89 for 100,000-499,999; n=52 for 500,000 or greater). A sample of 1000 tweets was obtained for each stratum and coded as being either about the organization or about personal-health topics. Subcategories for organization included information, engagement, and action. Subcategories for personal health included information and action.

Results: Of all LHD tweets (n=3000), 56.1% (1682/3000) related to personal health compared with 39.5% (1186/3000) that were about the organization. Of the personal-health tweets, 58.5% (984/1682) involved factual information and 41.4% (697/1682) encouraged action. Of the organization-related tweets, 51.9% (615/1186) represented one-way communication about the organization and its events and services, 35.0% (416/1186) tried to engage followers in conversation, and 13.3% (158/1186) encouraged action to benefit the organization (eg, attend events). Compared with large LHDs, small LHDs were more likely to post tweets about their organization (Cramer's V=0.06) but were less likely to acknowledge events and accomplishments of other organizations ($\chi^2=12.83$, $P=.02$, Cramer's V=0.18). Small LHDs were also less likely to post personal health-related tweets (Cramer's V=0.08) and were less likely to post tweets containing suggestions to take action to modify their lifestyle. While large LHDs were more likely to post engagement-related tweets about the organization (Cramer's V=0.12), they were less likely to ask followers to take action that would benefit the organization ($\chi^2=7.59$, $P=.02$, Cramer's V=0.08). While certain associations were statistically significant, the Cramer's V statistic revealed weak associations.

Conclusions: Twitter is being adopted by LHDs, but its primary use involves one-way communication on personal-health topics as well as organization-related information. There is also evidence that LHDs are starting to use Twitter to engage their audiences in conversations. As public health transitions to more dialogic conversation and engagement, Twitter's potential to help form partnerships with audiences and involve them as program participants may lead to action for improved health.

(*J Med Internet Res* 2013;15(8):e177) doi:[10.2196/jmir.2775](https://doi.org/10.2196/jmir.2775)

KEYWORDS

Twitter; social media; engagement; Internet; audience

Introduction

Organizational use of social media is expanding in both the private and public sectors [1,2]. Some evidence suggests social media is also being adopted in public health and health education settings [3-5]. Based on preliminary studies, inferences about the use of social media in public health settings include the following: (1) it is in an early adoption stage, (2) it tends to be used more often in urban and high-density populations compared with rural communities, and (3) it is used primarily to share information through one-way communication (ie, one sender and one or more receivers as with traditional mass media).

Some evidence suggests one-way communication delivered through social media can play an important role during disease outbreaks [6] and emergency response and relief efforts [7-9]. One-way communication can also create a foundation for more complex functions such as dialogue and mobilization [10]. However, limiting social media use to one-way communication decreases its interactive capacity to engage audiences. In fact, engaging audiences in two-way, dialogic, or conversational communication is the central purpose of social media [11]. So, while social media can be used to disseminate health information, it should also be used to create dialogue and engage audiences.

Engagement is a key element in mobilizing and building communities and the benefit of social media is not maximized unless it engages members of the community [10]. In the context of health promotion and social media, engagement has been defined as connections between people that contribute to a common good [12] and result in some type of action on behalf of the individual or organization [13]. This implies a mutual awareness and interaction between public health organizations and their audiences that lead to mutually beneficial outcomes. It has been recommended that social media be used more strategically within public health settings to engage audiences in ways that lead to action for health, such as involving them in the creation or delivery of programs or recruiting them as participants or recipients of programs, services, and activities [13].

One social media option for public health to engage audiences more effectively is Twitter, an information network composed of 140-character messages [14]. From 2010 to 2012, daily Twitter use experienced a four-fold increase; currently 15% of online adults use Twitter [15]. Corporate use of Twitter is also increasing significantly [16] and is used to enhance brands, increase visibility, support customers, network, communicate internally, generate leads, and support other online presences [17]. Twitter has re-launched its Twitter for Business site as an internal service for businesses who want to use Twitter to build communities and market and promote their products [18].

Nonprofit organizations are also using Twitter in a number of ways to promote their organizations and mobilize their audiences. In a study conducted among the 100 largest nonprofit organizations in the United States, Lovejoy and Saxton [10] examined the use of Twitter to determine the types of tweets (ie, messages) these organizations were sending to their audiences. Their message classification approach included three

categories: (1) information, or tweets about the organization (eg, highlights from events, news, facts, reports, etc), (2) community, which involved tweets that promoted dialogue and facilitated the creation of online communities, and (3) action, that involved tweets aimed at getting followers to do something for the organization. The researchers reported that 59% of tweets were informational in nature, 26% related to community building, and 15% related to action. They concluded that social media holds the potential to create opportunities for interpersonal engagement that are qualitatively different than traditional communication approaches and that to date, the nonprofit sector has overutilized the information function of social media and underutilized its interactive and dialogic components.

Research conducted among state health departments suggests Twitter is the most commonly used social media application in public health [3,4]. Initial analysis of county or local health departments (LHDs) revealed they are also using Twitter. According to the National Association of County and City Health Officials (NACCHO) [19], there are 2565 LHDs in the United States representing the largest delivery arm of public health services in the country.

LHD jurisdictions are classified as county (68%), multicounty (8%), and city (21%) with the remainder categorized as "other", including LHDs serving multiple cities and LHDs serving a county and a city not lying within the county boundaries [19]. However, LHD jurisdictions do not necessarily correlate with size of population served. For example, some county health departments serve small populations and others serve large populations and so forth. LHDs are granted legislative authority through codes and statutes and can be governed by local authorities (eg, local board of health or county or city elected officials) or by a state health agency or both [19].

The population-based primary prevention services provided most often by LHDs relate to chronic diseases and associated determinants. Clinical services provided most often by LHDs involve adult and child immunizations, infectious disease surveillance and screenings, food service inspections, and environmental health surveillance [19].

The LHD workforce constitutes a broad range of public health subdisciplines and professionals (eg, physicians, nurses, health educators, environmental health workers, emergency preparedness staff, nutritionists, etc) and thereby provides a potentially rich sample of Twitter use within public health [19]. However, to date, there is a paucity of literature about LHD use of Twitter, though one recent study did report how LHDs are using it to disseminate diabetes-related information [20]. Furthermore, while initial studies have reported frequency distributions of social media applications used in public health settings, no studies reported to date have investigated how social media is used within public health to engage audiences and involve them in actions related to programs and services.

Therefore, the purpose of this study was to examine how LHDs were using Twitter to communicate with and engage audiences. The following research questions guided this study:

1. Are LHDs more likely to use Twitter to share information, to engage with followers, or to promote action among followers?
2. Are there differences between LHDs serving varying population strata (eg, small, medium, and large) in the types of Twitter messages they post?
3. What health topics are LHDs tweeting about?

Methods

Sample

A list of all LHDs in the United States was obtained from the NACCHO (n=2565). Presence of a Twitter account was determined by three means. First, researchers visited the homepage of each LHD based on the website address provided by NACCHO. Researchers documented the presence or absence of a social media button indicating the LHD had a Twitter account. If there was no visual sign of Twitter on the homepage, a Google search was performed. Finally, on the Twitter homepage, each LHD name was entered in the search bar to confirm the presence of the account.

To be included in the study, the LHD had to have a Twitter account (n=306) and have posted a minimum of 50 tweets (n=210). Researchers then divided this list into three stratum: small, medium, and large based on size of population served (less than 100,000; 100-499,999; 500,000), hereafter referred to as small, medium, and large LHDs. Initial analysis showed a statistically significant difference in the total number of tweets a LHD posted based on size of population served, thereby confirming the need to stratify the study population. In the final sample, there were 69 small (32.8%), 89 medium (42.4%), and 52 large (24.8%) LHDs.

A complete list of tweets made by each LHD account was obtained using the Twitter Application Programming Interface (API) during July 2012. Because the Twitter API limits the maximum number of tweets that can be retrieved to the most recent 3200 per account, this limit was considered a complete tweet list for any LHD account that exceeded 3200 tweets. A total of 1000 tweets were sampled from each stratum for a total of 3000 tweets. To determine the number of tweets to be sampled from each LHD in the stratum, researchers divided 1000 by the number of LHDs in that stratum. There was an average of 14.5 tweets per LHD in the small stratum, 11.2 per LHD in the medium, and 19.2 tweets per LHD in the large stratum. Researchers randomly sampled tweets using a uniform distribution from each LHD account's list. This methodology was selected to reduce bias, including overrepresentation of tweet frequency during a specific event (eg, National Public Health Week), or underrepresentation due to irregularity of Twitter posts, or the LHD recently establishing a Twitter account.

Instrumentation

The researchers coded tweets based on the methodology used by Lovejoy and Saxton [10], though modified to reflect public health practice. The coding was designed to determine the purpose of the tweet including whether the LHD was using Twitter to disseminate information, foster a sense of community

or engagement with the community, or motivate audiences to action. Information sharing was defined as “one-way interaction, the exchange of information from the organization to the public” (page 343) [10]. Action was defined as “messages that aim to get followers to do something for the organization” (page 345) [10] or for their personal health. Based on social media terminology relevant in public health literature, the term “engagement” was used in lieu of “community” [13,21]. However, the researchers retained Lovejoy's and Saxton's definition of community for engagement, which was using Twitter to “interact, share, and converse with stakeholders in a way that ultimately facilitates the creation of an online community with its followers” (page 343) [10].

Tweets were first coded as being either about the organization or about personal-health topics. Then each tweet was coded for each of the categories described above (information, engagement, or action). Each classification category was mutually exclusive. Organization information included topics such as events or services provided, news, facts, reports, or job announcements. Organization engagement tweets were conversational in nature and gave thanks and recognition for doing something for the organization, acknowledged other organization's events, responded to public reply messages, asked for a response to a Twitter post, asked for feedback or suggestions, asked people to follow them or become a fan, or asked followers to spread the word or retweet the Twitter post. Action-based organization tweets invited followers to attend events, attend meetings and provide input, complete a survey, donate goods or money, volunteer time, or participate in lobbying or advocacy.

Personal-health tweets were limited to information and action. Information-related tweets involved general public health information (eg, eating foods rich in folic acid or taking a prenatal vitamin before you are pregnant can help prevent birth defects), risk communication related to disease outbreaks or natural or manmade disasters, or reports about public health conditions (eg, new report about obesity in America). Personal-health action-based tweets included messages to receive preventive health screenings (eg, get a mammogram), modify one's lifestyle (eg, make sure to walk for 30 minutes today), or to learn more and increase knowledge (eg, Now is the time to prepare for a disaster! Learn what you can do.). Each classification category was mutually exclusive.

The degree to which a tweet was considered interactive and attempted to engage the audience was determined by the presence of three components: (1) an @ reply symbol, signifying that the LHD was responding to a post made by another Twitter follower, (2) @username, indicating that the LHD was directing its post to a specific user, and (3) the use of personal pronouns [22,23]. The level of sophistication of each tweet was noted by (1) whether it was a truncated tweet, meaning that the LHD posted something on one platform (eg, Facebook) that was then posted to the Twitter account, (2) if it was a retweet, and (3) the existence of hashtags within the tweet. Truncated tweets and retweets denote that the LHD is not developing content specifically for Twitter but sharing what others have posted. Hashtags, which are used to categorize tweets so users can easily follow topics posted on Twitter, are also reflective of a more

advanced Twitter user. Tweets were also coded as to whether the follower was redirected to another source for more information. This signified that the LHD was using Twitter as a one-way communication tool and was linking people to more information.

Two research assistants pretested the coding instrument. Ten tweets from ten LHDs were selected for inclusion in the pretest analysis. Based on the results, minor adjustments were made to the coding sheet to clarify definitions. In some instances, the tweet content could be coded for more than one category. However, following Lovejoy and Saxton's methodology [10], research assistants identified the primary purpose of the tweet and coded it accordingly. If a personal-health tweet could be categorized as either information or action, it was coded as action with the assumption that taking action was the focus of the tweet. If a tweet included only a URL link with no other text, the tweet was viewed as only a redirect and was not coded further.

Four research assistants hand-coded all tweets. Two pairs of research assistants each coded half the tweets and then compared answers and resolved discrepancies. If there was a discrepancy between the two sets of coded data, the pair discussed each issue until consensus was reached. Discrepancies most often occurred because of a simple error related to data entry or a misinterpretation of the tweet.

Data Analysis

The chi-square test was used to assess differences between small, medium, and large LHDs. Standardized residuals were analyzed to determine which cells contributed significantly to the results. When a standardized residual is greater than 2, the cell is contributing significantly to the differences between groups [24]. The Cramer's V statistic was used to test the strength of association between two categorical variables. This is an appropriate test to use after the chi-square statistic is found to be significant [25].

Results

In the final sample, there were 69 small, 89 medium, and 52 large LHDs for a total of 210 LHD Twitter accounts. LHDs across all three strata (large, medium, and small) had a mean of 526.8 followers (ie, another Twitter user following the LHD) (range 9833, SD 1112.4) and followed a mean of 156.3 other Twitter users (range 4237, SD 342.99). Although large LHDs had more followers, followed more users, and posted more tweets compared with medium or small LHDs, these differences were not statistically significant as determined by the chi-square statistic (Table 1). The earliest Twitter account was established in June 2008 (Sacramento County, CA Public Health), and the most recent account was created in April 2012 (Independence, MO Health Department).

The majority of tweets (85.9%, $n=2578$) were original posts by the LHD. However, 15.2% (455/3000) of tweets were truncated, meaning the tweet originated from another source and was not fully displayed on Twitter. Just over 7% (7.3%, 218/3000) of posts were directly from Facebook. Almost 20% (19.7%, 592/3000) of tweets included the @ symbol, though this was

an @reply only 1% ($n=30$) of the time (ie, the LHD was responding to another Twitter user's post). Hashtags, used to mark keywords or topics in a tweet were used 16.2% (486/3000) of the time. Almost three-fourths of tweets (73.7%, 2211/3000) directed users to another source for more information through a website URL link (eg, the Centers for Disease Control and Prevention site). One third of tweets (36.7%, 1102/3000) used personal pronouns.

Overall, LHD tweets more often related to personal health (56.1%, 1682/3000) compared with information about the organization (39.5%, 1186/3000). There were a small number of posts that were non-health related and included topics such as news, events, and community happenings (2.2%, 65/3000).

Personal-Health Tweets

Of the personal-health tweets, the majority involved factual information (58.5%, 984/1682). Less than half of the personal-health tweets (41.4%, 697/1682) included an imperative verb that encouraged people to take action for their own health benefit. Of those tweets that included an action, 46.8% (326/697) related to knowledge (eg, learn more here), 37.2% (259/697) related to lifestyle behavior modification (eg, get more physical activity), and 10.4% (72/692) encouraged people to get preventive health screenings (eg, go for a mammogram).

Organization-Related Tweets

Organization-related tweets primarily represented one-way information about the organization (51.9%, 615/1186). The majority of these tweets (65.5%, 403/615) focused on events or services that the organization provided such as a flu clinic or breast-feeding class. Only 13% (13.3%, 158/1186) of organization-related tweets encouraged people to take action to benefit the organization. Of tweets aimed at getting people to take action, asking people to attend events was most common (69%, 109/158).

Just over a third (35%, 415/1186) of organization-related tweets were trying to engage their audiences in conversation. The most common way to engage with audiences was through acknowledgement of other organizations' events (23.9%, 99/415) and giving thanks for recognition (17%, 70/415). In addition, several tweets appeared conversational in nature (24.1%, 100/415), while not fitting the exact engagement categories identified in the coding sheet (eg, "Back from a refreshing but hot walk. Now following heat safety guidelines and drinking some water #heatwave").

Tweets by Size of Population Served

There were significant, yet weak associations between size of population served by the LHD and several of the study variables (Table 2). Analysis of standardized residuals showed that either the large or small LHDs were contributing to the significant chi-square value, but never the medium-sized LHDs. Small LHDs were more likely to have truncated tweets (Cramer's $V=0.18$) and use personal pronouns (Cramer's $V=0.06$). Large LHDs were more likely to use the @ symbol in their original tweets (Cramer's $V=0.09$), redirect followers to a different

website or link (Cramer's $V=0.096$), and use hashtags (Cramer's $V=0.14$)

There was a significant association between LHD size and posting personal-health-related tweets. Examination of standardized residuals showed that small LHDs were less likely to do so (Cramer's $V=0.08$). For personal-health-related tweets, small LHDs were less likely to post tweets containing suggestions to take action to modify their lifestyle ($\chi^2=8.90$, $P=.01$, Cramer's $V=0.11$).

Overall, small LHDs were more likely to post tweets about the organization (Cramer's $V=0.06$), and large LHDs were more likely to post organization-engagement-related tweets (Cramer's $V=0.12$). For organization-engagement tweets, small LHDs were least likely to acknowledge events and accomplishments of other organizations ($\chi^2=12.83$, $P=.02$, Cramer's $V=0.18$). Large LHDs were also less likely to ask followers to take action that would benefit the organization ($\chi^2=7.59$, $P=.02$, Cramer's $V=0.08$).

Table 1. Number of tweets, followers, and who they follow by local health department size.

	Small n=69	Medium n=89	Large n=52	Total n=210	P value
Mean number of lifetime tweets (SD)	334.1 (959.8)	443.5 (620.9)	748.7 (899.6)	483.1 (827.7)	.31
Mean number of users they are following (SD)	266.1 (430.2)	147.0 (195.7)	293.5 (616.3)	156.3 (342.99)	.14
Mean number of followers (SD)	64.78 (64.5)	369.1 (534.6)	1341.8 (1894.93)	526.8 (1112.4)	.40

Table 2. Differences between small, medium, and large LHDs reported by frequency and percentage.

Variable	Small	Medium	Large	Total	P value
Original tweet	851 (33%)	871 (87.1%)	856 (86.6%)	2578/3000 (85.9%)	.41
Tweet truncated	231 (50.8%)	152 (15.2%)	72 (7%)	455/3000 (15.2%)	<.001
Tweet truncated from Facebook	109 (50%)	82 (38%)	27 (12%)	218/3000 (7.3%)	<.001
Retweet	147 (35.3%)	130 (31.2%)	140 (33.6%)	417/3000 (13.9%)	.54
@Reply	3 (10%)	8 (27%)	19 (63%)	30/3000 (0.01%)	.001
@username	178 (30.1%)	184 (31.1%)	230 (38.9%)	592/3000 (19.7%)	.06
Redirect to another link, websites, etc, to learn more	706 (31.9%)	708 (32.0%)	797 (36%)	2211/3000 (73.7%)	<.001
Use of personal pronouns (1st and 2nd person)	401 (36.4%)	369 (33.5%)	332 (30.1%)	1102/3000 (36.7%)	.006
Hashtags (#)	116 (23.9%)	135 (27.8%)	235 (48.4%)	486/3000 (16.2%)	<.001
Organization—Overall	438 (36.9%)	382 (32.2%)	366 (30.9%)	1186/3000 (39.5%)	.003
Organization—Information	241 (39.2%)	201 (32.7%)	173 (28.1%)	615/1186 (51.9%)	.08
Organization—Engagement	132 (31.8%)	125 (30.1%)	158(38.1%)	415/1186 (35.0%)	<.001
Organization—Action	68 (43%)	56 (35%)	34 (22%)	158/1186 (13.3%)	.02
Personal Health—Overall	509 (30.3%)	574 (34.1%)	599 (35.6%)	1682/3000 (56.0%)	<.001
Personal Health—Information	322 (32.7%)	321 (32.6%)	341 (34.7%)	984/1682 (58.5%)	.03
Personal Health—Action	187 (26.8%)	254 (36.4%)	256 (36.7%)	697/1682 (41.4%)	.03

Discussion

Principal Findings

This study examined how LHDs are using Twitter to communicate with their followers. Results showed that LHDs are more likely to use Twitter to convey personal-health information compared with information about the organization. Personal-health tweets were more often factually based with encouragement to take action to learn more. Organization-based tweets were generally related to events and services LHDs provided with an invitation to attend and support the events. There were some differences between large and small LHDs.

Nearly 12% of LHDs use Twitter, which is consistent with an earlier report of 13% from NACCHO [19]. This is also fairly similar to individual use of Twitter reported at 15% [15], but lower than use of Twitter within nonprofit organizations [10], large companies [26], and state health departments [3,4]. One explanation for lower use of Twitter among LHDs compared with nonprofit organizations may relate to revenue and funding streams. Nah and Saxton [27] reported that organizations that rely on donor-based funding were more likely to use social media than those funded by the government. Although LHDs appear to be using Twitter to help fulfill their public health mission, their funding, though often tenuous, is not influenced by relationships with their Twitter followers.

In the case of lower Twitter use among LHDs compared with state health departments, it is not uncommon for state health departments to employ a larger and more diverse workforce associated with broader capacity. Whereas a state health department might have the capacity to designate a staff member as a public information officer or social media specialist, this is less likely to exist among LHDs, particularly small, understaffed LHDs funded primarily by Medicaid and Medicare with mandates to provide a range of clinical services [19].

Twitter, like other social media applications, can help public health strategically establish brands, foster relationships with consumers, and promote its organizations as well as their products and services. However, LHDs are using Twitter primarily as a way to distribute personal-health information, which is inconsistent with ways many other organizations use Twitter. While some LHDs are using Twitter for organization-based purposes, this core element and function of Twitter is clearly underutilized.

LHDs' predominant use of Twitter to share information does relate to one of the ten essential public health services to inform, educate, and empower people [28]. This may be viewed by public health practitioners as a general mandate that can be addressed through a social media application like Twitter and may explain why LHDs post more information about personal health on Twitter compared with information about their organizations. However, there is little evidence that using Twitter as a one-way communication tool to disseminate health information is effective at improving health status.

While some evidence suggests that broad dissemination of information characterized by traditional mass media campaigns can improve population health, effective campaigns require

simultaneous availability of and access to programs, services, and products that facilitate change [29]. Furthermore, broad dissemination of information ignores the fact that messages should be targeted to the intended audience. In the case of Twitter, LHDs may know nothing or very little about their followers unless they engage in dialogic communication to establish relationships. To indiscriminately post information on Twitter is inefficient. In fact, this contributes to what has been described as a fractured and cluttered media environment that can be resolved only through careful planning and testing of campaign content with intended audiences [29].

It was encouraging that at least one-third of LHD tweets attempted to engage followers, foster relationships, create networks, or build communities. These results are similar to those found by Lovejoy and Saxton in their analysis of how nonprofit organizations use social media [10]. Use of personal pronouns was present over a third of the time and more common among smaller LHDs. Additionally, evidence of effort toward dialogic communication included tweets that tended to be conversational in nature and may have used personal pronouns but were not necessarily intended for the purpose of engagement. This evidence of more conversational posting indicates LHDs may be trying to create a Twitter persona that is warm and friendly, thus making it more inviting for Twitter users to follow.

Data suggesting LHDs are trying to engage with audiences has not been reported previously. While it has been reported that state health departments are using Twitter almost exclusively for one-way communication [3,4], research indicates that when an organization's communication is more interactive, the result is a better relationship with its consumers [30]. In turn, better relationships with consumers can lead to higher levels of engagement. As reported by Neiger et al [13] in their three-phase engagement hierarchy, use of social media should culminate in high engagement characterized by online or offline audience member involvement with the organization's programs or services either as a partner or a participant.

There were a few differences among small and large LHDs in terms of tweet composition or content. While the strength of these associations is small, the relationships are instructive. First, large LHDs appear to be more sophisticated in the technical use of Twitter as evidenced by using hashtags more frequently. Both large and small LHDs appear to struggle with developing original content. For example, small LHDs are more likely to post truncated tweets (ie, they were posted somewhere else first such as Facebook then later appeared on Twitter). Large LHDs tended to redirect followers to other sites for more information. This indicates LHDs may lack either the technical capacity or general commitment to create original content on Twitter that more effectively develops relationships and engages followers.

The content of tweets also varied among LHDs based on size of population served. Small LHDs were more likely to post about the organization and less likely to post personal-health tweets. Since small LHDs may have less organizational capacity and may be focused on a more finite set of clinical services [19], they may be less inclined to disseminate personal-health information unrelated to their services or to attempt to modify

the lifestyles of their followers. Small LHDs posting about their organization indicates they may be more interested in personal relationships with their clients and becoming acquainted with and connecting with their audiences.

In their organization-engagement tweets, small LHDs were less likely to acknowledge the activities of other organizations. Since small LHDs serve less densely populated areas and are typically located in rural, more isolated locations, it is reasonable they may be less likely to acknowledge other organizations that are more physically inaccessible and removed from their own clients.

Small LHDs more often asked followers to do something for the organization. This may be due to a limited capacity of small LHDs to provide a wide range of services or it may relate to a sense of familiarity or cohesion that might be more common within rural communities that are served by smaller LHDs. A long-held belief in the delivery of mental health services in rural communities, represented primarily by smaller LHDs in this study, are more closely knit than urban communities [31].

Limitations

Results should be interpreted with the following limitations in mind. First, to be included in this study LHDs had to have a Twitter account with a minimum of 50 tweets. There may be more LHDs now that would qualify for such a study. In addition, a cross-sectional survey of LHD tweets was obtained using the Twitter API during July 2012. The prevalence and type of LHD Twitter use may have changed somewhat from that point to the present. Also, in distinguishing between the primary purpose and general content of tweets (ie, information, engagement or action), coder subjectivity was a limitation. However, coders compared interpretations and resolved discrepancies thereby increasing intercoder reliability. Finally, while certain associations of data were found to be statistically significant, the Cramer's V statistic revealed weak associations.

Conclusions

Twitter is being adopted by LHDs, but its primary use involves disseminating one-way information on personal-health topics as well as organization-related information. There is also evidence that LHDs are starting to use Twitter to engage their audiences in conversations.

Since a paucity of evidence supports the use of Twitter or other forms of social media to disseminate one-way information as

a stand-alone intervention to improve health status, LHDs should transition to more dialogic communication. More specifically, LHDs should use Twitter to develop relationships with their followers (ie, individuals and organizations) to create partnerships that leverage resources and also increase participation in LHD programs and services with the intent of improving health status [13]. In using Twitter to develop relationships, LHDs should post more original content including information about their organizations. Conversely, LHDs should post fewer truncated tweets and redirect followers less often to other sites for information.

LHDs should also develop strategic implementation and communication plans that include forethought of how Twitter or other forms of social media could be integrated and used most effectively. For example, if Twitter is used to engage audiences and increase partnerships and program participation, strategies must also be in place to activate and sustain partnerships and program participation.

If strategic communication plans identify that priority audiences prefer Twitter as a communication channel, then Twitter should be used more effectively to reach the intended audience rather than being used indiscriminately. If communication plans do not suggest that members of priority audiences have access to social media applications or they are not preferred communication channels, other forms of communication will be more appropriate.

This study has helped identify initial patterns of Twitter use among LHDs. Future research should include investigations that help determine why LHDs actually use Twitter or other forms of social media. These studies could further examine perceived benefits of engagement or the relationship between engagement and partnership and participation outcomes. Related outcomes may be of particular interest to public health funding agencies that support social media research.

With the increasing popularity of social media, public health has unprecedented opportunities to communicate directly with its audiences. As public health more fully uses social media to engage these audiences and further research clarifies how this can be done most effectively, the potential of social media to aid in change efforts that improve health status will be better understood and applied.

Acknowledgments

The authors wish to acknowledge NACCHO and its National Profile of Local Health Departments and for all data related to local health departments. The authors also acknowledge Emily Christensen and Victoria Ducette for their work in organizing and coding data, Dr Christophe G Giraud-Carrier for conceptual and technical assistance, and Mark and Christine Cronquist for financial assistance to hire research assistants.

Conflicts of Interest

None declared.

References

1. Barnes N, Lescault A, Andonian J. Social Media Surge by the Fortune 500: Increase use of Blogs, Facebook, Twitter and more. 2012. URL: <http://www.umassd.edu/cm/socialmedia/2012fortune500/> [accessed 2013-06-06] [WebCite Cache ID 6HBH5fd7d]
2. Stelzner M. Social Media Marketing Industry Report. 2012. URL: <http://www.socialmediaexaminer.com/SocialMediaMarketingIndustryReport2012.pdf> [accessed 2013-06-06] [WebCite Cache ID 6HBOU8iDB]
3. 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](https://doi.org/10.1186/1471-2458-12-242)] [Medline: [22449137](https://pubmed.ncbi.nlm.nih.gov/22449137/)]
4. Harris JK, Snider D, Mueller N. Social media adoption in health departments nationwide: The state of the states. *Frontiers in Public Health Serv and Syst Res* 2013;2(1):---
5. Hanson C, West J, Neiger B, Thackeray R, Barnes M, McIntyre E. Use and Acceptance of Social Media Among Health Educators. *American Journal of Health Education* 2011 Jul;42(4):197-204. [doi: [10.1080/19325037.2011.10599188](https://doi.org/10.1080/19325037.2011.10599188)]
6. Centers for Disease Control and Prevention. CDC Flu. 2013. URL: <https://twitter.com/CDCFlu> [accessed 2013-06-06] [WebCite Cache ID 6HBIDJqcf]
7. Merchant RM, Elmer S, Lurie N. Integrating social media into emergency-preparedness efforts. *N Engl J Med* 2011 Jul 28;365(4):289-291. [doi: [10.1056/NEJMp1103591](https://doi.org/10.1056/NEJMp1103591)] [Medline: [21793742](https://pubmed.ncbi.nlm.nih.gov/21793742/)]
8. Smith BG. Socially distributing public relations: Twitter, Haiti, and interactivity in social media. *Public Relations Review* 2010 Nov;36(4):329-335. [doi: [10.1016/j.pubrev.2010.08.005](https://doi.org/10.1016/j.pubrev.2010.08.005)]
9. Hughes AL, Palen L. Twitter adoption and use in mass convergence and emergency events. *IJEM* 2009;6(3/4):248-360. [doi: [10.1504/IJEM.2009.031564](https://doi.org/10.1504/IJEM.2009.031564)]
10. Lovejoy K, Saxton GD. Information, community and action: How nonprofit organizations use social media. *J Comput-Mediated Commun* 2012;17:337-353. [doi: [10.1111/j.1083-6101.2012.01576.x](https://doi.org/10.1111/j.1083-6101.2012.01576.x)]
11. Safko L, Blake DK. *The Social Media Bible: Tactics, Tools and Strategies for Business Success*. Hoboken, NJ: John Wiley; 2009.
12. 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/)]
13. 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/)]
14. Twitter. The Twitter Glossary. 2013. URL: <https://support.twitter.com/groups/31-twitter-basics/topics/104-welcome-to-twitter-support/articles/166337-the-twitter-glossary> [accessed 2013-06-06] [WebCite Cache ID 6HBIGpEW9]
15. Smith A, Brenner J. *Twitter Use*. Washington, DC: Pew Internet & Life Project; 2012. URL: <http://www.pewinternet.org/Reports/2012/Twitter-Use-2012.aspx> [accessed 2013-06-06] [WebCite Cache ID 6HBJ2Fyg0]
16. Shane D. Small Business Trends: Company Use of Twitter is Significantly Increasing. 2012. URL: <http://smallbiztrends.com/2012/09/company-use-of-twitter-is-significantly-increasing.html> [accessed 2013-06-06] [WebCite Cache ID 6HBJADkH5]
17. Dugan L. Seven ways businesses can use Twitter. 2012. URL: http://www.mediabistro.com/alltwitter/7-ways-businesses-can-use-twitter_b3994 [accessed 2013-06-06] [WebCite Cache ID 6HBJHOPAE]
18. Sophy J. Twitter Relaunches its Twitter for Business Site. 2013. URL: <http://smallbiztrends.com/2013/04/twitter-for-business-relaunched.html> [accessed 2013-06-06] [WebCite Cache ID 6HBJR0E5F]
19. National Association of County and City Health Officials. *National Profile of Local Health Departments*. 2010. URL: http://www.naccho.org/topics/infrastructure/profile/resources/2010report/upload/2010_Profile_main_report-web.pdf [accessed 2013-06-06] [WebCite Cache ID 6HBMR1qOi]
20. 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](https://doi.org/10.5888/pcd10.120215)] [Medline: [23639765](https://pubmed.ncbi.nlm.nih.gov/23639765/)]
21. #SMM Standards. About SMM Standards. 2012. URL: <http://www.smmstandards.com/> [accessed 2013-06-06] [WebCite Cache ID 6HBN3krqE]
22. Stewart S. *Social Media Strategies for Small Business*. 2012. URL: <http://www.metzger.com/social-media-strategies-for-small-business/> [accessed 2013-06-06] [WebCite Cache ID 6HBNRpsmK]
23. Dann S. Twitter content classification. *First Monday* 2010;15(12):- [doi: [10.5210/fm.v15i12.2745](https://doi.org/10.5210/fm.v15i12.2745)]
24. Hinkle DE, Wiersma W, Jurs, SG. *Applied statistics for the behavioral sciences*. Boston: Houghton Mifflin; 1988.
25. Pett MA. *Nonparametric statistics for health care research: statistics for small samples and unusual distributions*. Thousand Oaks, Calif: Sage Publications; 1997.
26. Culnan MJ, McHugh PJ, Zubillaga JI. How large US companies can use Twitter and other social media to gain business value. *MIS Q Executive* 2010;9(4):243-259.
27. Nah S, Saxton GD. Modeling the adoption and use of social media by nonprofit organizations. *New Media & Society* 2012 Jul 31;15(2):294-313. [doi: [10.1177/1461444812452411](https://doi.org/10.1177/1461444812452411)]
28. American Public Health Association. *10 Essential Public Health Services*. 2013. URL: <http://www.apha.org/programs/standards/performancestandardsprogram/resexentialservices.htm> [accessed 2013-06-06] [WebCite Cache ID 6HBNkZIFN]

29. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. *Lancet* 2010 Oct 9;376(9748):1261-1271. [doi: [10.1016/S0140-6736\(10\)60809-4](https://doi.org/10.1016/S0140-6736(10)60809-4)] [Medline: [20933263](https://pubmed.ncbi.nlm.nih.gov/20933263/)]
30. Saffer AJ, Sommerfeldt EJ, Taylor M. The effects of organizational Twitter interactivity on organization-public relations. *Public Relations Rev* 2013;39(3):213-215.
31. Blank MB, Fox JC, Hargrove DS, Turner JT. Critical issues in reforming rural mental health service delivery. *Community Ment Health J* 1995 Dec;31(6):511-524. [Medline: [8608697](https://pubmed.ncbi.nlm.nih.gov/8608697/)]

Abbreviations

API: Twitter Application Programming Interface

LHD: local health department

NACHHO: National Association of County and City Health Officials

Edited by G Eysenbach; submitted 14.06.13; peer-reviewed by K Lovejoy, M Laurent; comments to author 05.07.13; revised version received 26.07.13; accepted 04.08.13; published 19.08.13.

Please cite as:

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

URL: <http://www.jmir.org/2013/8/e177/>

doi: [10.2196/jmir.2775](https://doi.org/10.2196/jmir.2775)

PMID: [23958635](https://pubmed.ncbi.nlm.nih.gov/23958635/)

©Brad L Neiger, Rosemary Thackeray, Scott H Burton, Callie R Thackeray, Jennifer H Reese. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 19.08.2013. 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.

Review

The Impact of Social Media on Medical Professionalism: A Systematic Qualitative Review of Challenges and Opportunities

Fatemeh Gholami-Kordkheili¹, DDS; Verina Wild², MD; Daniel Strech¹, MD, PhD

¹Institute for History, Ethics and Philosophy of Medicine, CELLS-Centre for Ethics and Law in the Life Science, Hannover Medical School, Hannover, Germany

²Institute for Biomedical Ethics, University of Zurich, Zurich, Switzerland

Corresponding Author:

Daniel Strech, MD, PhD

Institute for History, Ethics and Philosophy of Medicine

CELLS-Centre for Ethics and Law in the Life Science

Hannover Medical School

Carl-Neuberg-Street 1

Hannover, 30625

Germany

Phone: 49 511 532 2709

Fax: 49 511 532 5650

Email: strech.daniel@mh-hannover.de

Abstract

Background: The rising impact of social media on the private and working lives of health care professionals has made researchers and health care institutions study and rethink the concept and content of medical professionalism in the digital age. In the last decade, several specific policies, original research studies, and comments have been published on the responsible use of social media by health care professionals. However, there is no systematic literature review that analyzes the full spectrum of (1) social media-related challenges imposed on medical professionalism and (2) social media-related opportunities to both undermine and improve medical professionalism.

Objective: The aim of this systematic qualitative review is to present this full spectrum of social media-related challenges and opportunities.

Methods: We performed a systematic literature search in PubMed (restricted to English and German literature published between 2002 and 2011) for papers that address social media-related challenges and opportunities for medical professionalism. To operationalize “medical professionalism”, we refer to the 10 commitments presented in the physicians’ charter “Medical professionalism in the new millennium” published by the ABIM Foundation. We applied qualitative text analysis to categorize the spectrum of social media-related challenges and opportunities for medical professionalism.

Results: The literature review retrieved 108 references, consisting of 46 original research studies and 62 commentaries, editorials, or opinion papers. All references together mentioned a spectrum of 23 broad and 12 further-specified, narrow categories for social media-related opportunities (n=10) and challenges (n=13) for medical professionalism, grouped under the 10 commitments of the physicians’ charter.

Conclusions: The accommodation of the traditional core values of medicine to the characteristics of social media presents opportunities as well as challenges for medical professionalism. As a profession that is entitled to self-regulation, health care professionals should proactively approach these challenges and seize the opportunities. There should be room to foster interprofessional and intergenerational dialogue (and eventually guidelines and policies) on both challenges and opportunities of social media in modern health care. This review builds a unique source of information that can inform further research and policy development in this regard.

(*J Med Internet Res* 2013;15(8):e184) doi:[10.2196/jmir.2708](https://doi.org/10.2196/jmir.2708)

KEYWORDS

social media; professionalism; facebook; blogs; Twitter; health policy

Introduction

Professionalism is the basis of medicine's contract with society [1]

In 2002, the European Federation of Internal Medicine, the American College of Physicians-American Society of Internal Medicine (ACP-ASIM), and the American Board of Internal Medicine (ABIM) felt it necessary to renew the sense of professionalism due to changing market forces. The result of these efforts was a new physicians' charter, which claimed to apply to physicians throughout the world.

Ten years later, the rising influence of social media in our private and professional lives is a new force that affects our understanding of medical professionalism. Social media, as a part of the Web 2.0, include blogs, wikis, podcasts, and social networking platforms such as Twitter, LinkedIn, YouTube, and Facebook, to name just a few. In contrast to websites where people are limited to the passive viewing of content, Web 2.0 tools are people-based knowledge sharing, learning, social interaction, and collective intelligence tools that support knowledge collaboration, exchange, sharing, and creation [2]. Thompson et al reported in 2008 that 45% of medical trainees, 64% of medical students, and 13% of medical residents had Facebook accounts [3].

The asymmetry of disclosure in the doctor-patient relationship was emphasized long before social media [4]. Today, social media allow patients to gather increasingly more information about their doctors' private and professional life. Excessive self-disclosure from the side of the physician is generally regarded as a boundary violation in the patient-physician treatment relationship [5]. Disclosure of this kind of personal information on a social networking site is usually not aimed at patients, but patients might nevertheless access this information [6].

Persistence, searchability, replicability, and invisible audiences are unique characteristics of Facebook and other social media platforms [7], which form—based on the ease of searching and storing digital information—a “permanent” digital fingerprint and online reputation. Once information is online, it is extremely difficult to remove it (if at all) and it can quickly spread beyond one's control. A moment of rashness could have unintended and irreversible consequences in the future such as suspension from medical school, loss of employment as a physician, and loss of trust in the medical profession [8]. It could concern future or current employment candidacy, or current employment and training conditions. There are already cases of students, trainees, or medical staff being dismissed because of their “unprofessional” online image [9,10].

However, the reduction of power imbalances between patients and doctors has been shown to improve patient confidence in starting, stopping, or making changes to treatment regimens [11]. Social media may also help to distribute precise health information to a larger group of individuals than ever before.

But is online available medical information reliable? Who provides the medical information on blogs, YouTube, Twitter, and Facebook? In 2008, there were 1434 medical-related blogs; however, only 279 were actually written by medical professionals [12]. As advertising and business interests strongly influence the order of search engine listings [13], it might be advisable for the medical and dental professions to proactively refer patients to high-quality sources of medical online information [14,15].

Universities and medical organizations, especially in the United States (such as the American Medical Association, AMA) and United Kingdom, have started to develop guidelines and policies for health care professionals concerning proper social media use. In order to foster awareness, courses on handling social media associated with medical professionalism have been implemented in the professional curricula [16]. The recently published position paper on online medical professionalism by the American College of Physicians and the Federation of State Medical Boards provides the latest recommendations on strategies for physician-physician communication that aims at preserving confidentiality while best profiting from the new technologies of social media [17].

The importance of social media is also indicated by the increasing number of scientific publications that deal with them in the medical context. While our search (see Methods) found a total of 1471 publications focusing on social media on PubMed in December 2011, by the end of December 2012 there were 2330 hits.

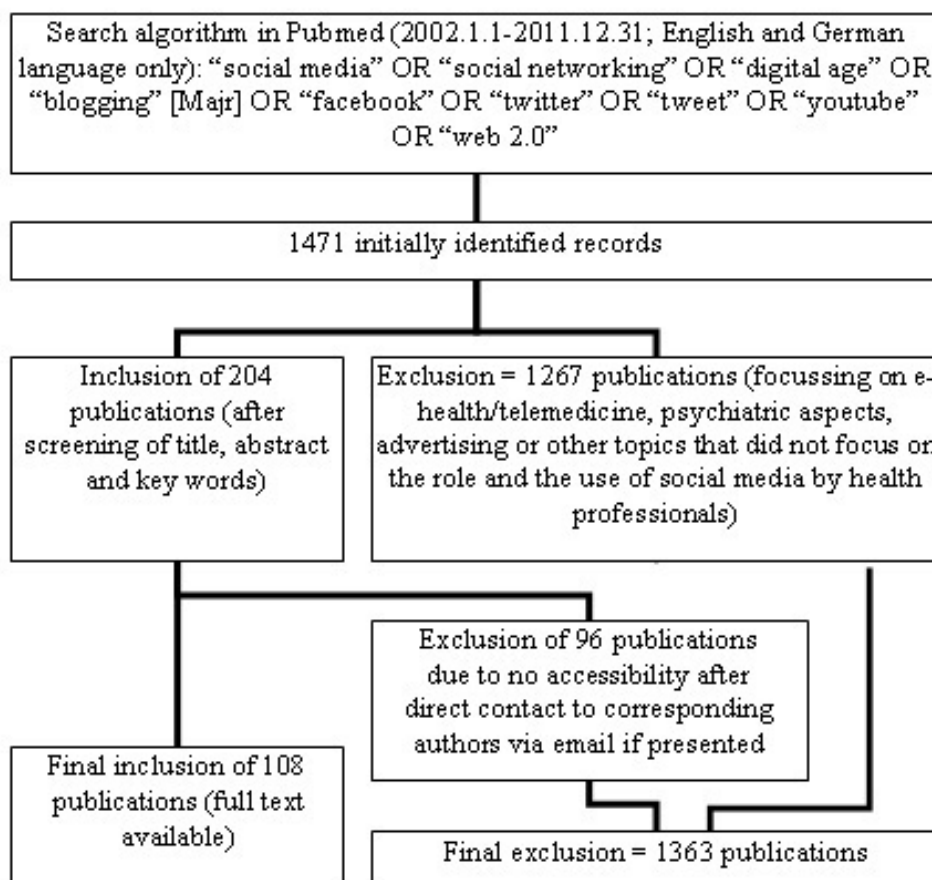
To our knowledge, there is no systematic literature review that analyses the full spectrum of (1) social media-related challenges to medical professionalism and (2) social media-related opportunities to either undermine or improve medical professionalism. The aim of this systematic qualitative review is to present this spectrum.

Methods

Literature Search and Eligibility Criteria

In December 2011, we searched PubMed with the following terms: “social media” OR “social networking” OR “digital age” OR “blogging” [Majr] OR “facebook” OR “twitter” OR “tweet” OR “youtube” OR “Web 2.0”. The search was restricted to English or German language papers. Publications before 2002 were excluded because all major social media platforms were founded after 2002: MySpace was founded in 2003 [18], Facebook in 2004 [19], and Twitter in 2006 [20]. We included publications focusing on the use of social media by health professionals, challenges imposed on health professionals by social media use, and ethical considerations concerning the relationship between patients and health professionals in the Internet era. We excluded publications focusing on eHealth/telemedicine, addiction, and other psychiatric issues related to social media, and advertising or marketing. See Figure 1.

Figure 1. Flowchart illustrating identified references.



Extraction and Categorization of Social Media–Related Opportunities and Challenges for Medical Professionalism

Our aim was to develop a qualitative framework of narrow and broad categories of social media–related opportunities and challenges for medical professionalism that best accommodated the opportunities and challenges mentioned in the included publications.

To operationalize “medical professionalism”, we referred to the 10 commitments/professional responsibilities presented in the physicians’ charter, “Medical professionalism in the new millennium” published by the ABIM Foundation, the ACP-ASIM Foundation, and the European Federation of Internal Medicine. To our knowledge, the physicians’ charter is the most widely accepted and most often cited framework for medical professionalism. It has been endorsed by over 90 professional societies worldwide. Since its publication in 2002 in several journals, it has been cited more than 900 times (as assessed by Scopus). The 10 commitments are (1) professional competence, (2) honesty with patients, (3) patient confidentiality, (4) maintaining appropriate relations with patients, (5) improving quality of care, (6) improving access to care, (7) a just distribution of finite resources, (8) scientific knowledge, (9)

maintaining trust by managing conflicts of interest, and (10) professional responsibilities.

We employed the 10 commitments of medical professionalism as our matrix to guide the identification of text passages that mention social media–related opportunities or challenges for medical professionalism. Mentions of such opportunities and challenges in different papers were compared. Broad and narrow categories were developed for similar mentions of opportunities and challenges. According to our matrix, these broad and narrow categories were grouped under the 10 commitments.

To ensure the validity of coding as well as intercoder reliability, we employed the following procedure: 3 authors (FG, VW, DS) identified and initially categorized opportunities and challenges (based on the above described extraction matrix) independently in a subsample of 5 publications. The authors discussed whether paragraphs mentioned opportunities and challenges and how they should be categorized. The remaining 103 publications were grouped in three clusters of 60, 23, and 20 publications. One author (FG) with an MD degree then extracted and categorized social media-related opportunities and challenges from this first cluster of publications. The result was a first version of the spectrum of social media-related opportunities and challenges grouped under the 10 commitments. The second and third clusters of references were then used to check theoretical saturation of the spectrum. Theoretical saturation

means that no new categories can be generated [21]. Once theoretical saturation was reached for broad categories, the other authors (DS, VW), with professional backgrounds in bioethics, clinical psychiatry, internal medicine, philosophy, and health services research, checked the extraction and categorization of opportunities and challenges in a random sample of 25 publications. Coding problems were resolved by frequent meetings and discussions between all authors.

Results

From 1471 initial hits in PubMed, we finally included 108 in this review. The 108 references consist of 46 original research studies and 62 commentaries, editorials, and opinion papers. The majority are from the United States (79 publications), followed by 15 from the United Kingdom. Other papers come from Canada (5 publications), Ireland (3 publications), Australia (2 publications), and Germany, Peru, France, and New Zealand (1 publication each). The sample consists of one article published in 2006, three in 2008, 13 in 2009, 21 in 2010, and 70 articles in 2011.

We identified 23 broad and 12 further-specified narrow categories for social media–related opportunities ($n=10$) and challenges ($n=13$) for medical professionalism, grouped under the 10 commitments of the physicians' charter.

For example, for the first commitment “professional competence”, we identified four broad categories for opportunities (A-D) and one broad category for a challenge (E): (A) Employing social media as a tool for improved information sharing, (B) Increasing the involvement by doctors in under-served areas, (C) Committing to life-long learning supported by the use of social media, (D) Mentoring student's reasonable engagement in social media, and (E) Ensuring evidence-based Continuing Medical Education in the environment of social media. Some of these broad categories are specialized into more narrow categories. For example, the broad category (A) Employing Social Media as a tool for improved information sharing was specified into five narrow categories: (A1) Fast and boundless dissemination of news and experience, (A2) Collaboration on challenging cases, (A3) Improving access to and benefits of conferences and news exchange, (A4) Sharing information on physician-only social media sites, and (A5) Accessing news/information from professional organizations. One of many original text passages extracted from the narrow category (A1) is “With Internet-based tools, physicians are no longer limited by geography, specialty, and time zone in their attempts to connect, engage, and learn from each other” [22]. For technical reasons and for didactic purposes, we restrict our presentation to one exemplary text passage for each of the 33 narrow categories (see [Multimedia Appendix 1](#) for these findings; [4,9,14,22-41]).

Discussion

Principal Findings

This systematic qualitative review presents the full spectrum of social media–related opportunities and challenges for medical professionalism as they are currently discussed in original

research studies, commentaries, editorials, or opinion papers published in scientific journals listed in PubMed. Thereby it builds a unique source of knowledge that can inform further research and policy development in the intersection of social media and medical professionalism.

The need for policies on the use of social media by medical professionals, trainees, and students has already been addressed by some universities [42] and also by institutions such as the AMA [43]. The AMA policy “Medical professionalism in the digital age”, which was adopted in November 2010, presents general recommendations. It encourages the medical practitioner to “weigh a number of considerations” when it comes to social media. The gist of the policy is to preserve patient privacy and confidentiality in all environments, to avoid excessive self-disclosure by using adequate privacy settings, being aware that they are not absolute, and routinely monitoring one's online presence. It stresses the necessity of maintaining appropriate patient/physician boundaries, and in doing so to consider the separation of professional and personal online content. The policy tries to raise awareness of the professional's responsibility to bring posted unprofessional content to the attention of the individual in question or to inform appropriate authorities, as those failures may affect the medical professional's reputation among patients and colleagues and may undermine public trust. Even though the above-mentioned issues (which almost all describe challenges) are important, the AMA policy neither illustrates a more differentiated view of social media-related challenges, nor does it acknowledge social media-related opportunities and the need to address them appropriately. Such opportunities include, for instance, improvement in sharing information, access to care, and quality of care, etc [43] (see [Multimedia Appendix 1](#)).

The University of Florida, for example, recognizes the relevance of social media as a current form of communication. However, it also focuses on challenges and distinguishes “strictly forbidden” from “strongly discouraged” online interactions, which could be the basis for disciplinary actions. Violating patient confidentiality, reporting private academic information, and neglecting official work commitments when interacting online are strictly forbidden actions. Strongly discouraged actions include use of vulgar language, implying disrespect for any individual due to age, race, gender, etc, presentation of alcohol misuse, substance abuse, sexual promiscuity, and posting unflattering material on another individual's website. The policy tries to raise awareness that a mature, responsible, and professional attitude should also be displayed when interacting online privately and to think twice before posting any material because online privacy measures might be unreliable [44].

Although it is a laudable first step that both the AMA policy and the University of Florida policy explicitly address some social media–related challenges for medical professionalism, in their current version they address neither the full spectrum of challenges nor any of the social media–related opportunities (see [Multimedia Appendix 1](#)). In general, social media–related challenges are more frequently discussed in the reviewed publications than social media–related opportunities. But as the relevance of social media might further increase, there is an ongoing demand for a critical and constructive discussion about,

and guidelines/recommendations on, how to best possibly address the multifaceted spectrum of challenges and opportunities.

Particularly among medical students and young professionals on the one hand and educators and practicing physicians on the other, there may be a different attitude towards the use of social media. Prensky introduced the distinction of digital natives and digital immigrants that is often referred to in today's debate on online medical professionalism [23,45]. Current trainees and medical students born after 1980 are considered as digital natives, as they grew up in a world where using technology (eg, computers, the Internet, text messaging, blogging, and SMS text messaging) was already integrated within their education, patterns of establishing/maintaining relationships, and means of self-expression. Older faculty who completed their training before 1980 are considered digital immigrants because a good number of them experience a challenge to continually adopt to the particularities of the digital age with which their students are likely more familiar [23]. However, a sharp distinction between digital natives and digital immigrants might blur in the near future, and further distinctions across digital natives might occur. We have, for example, anecdotal evidence that some current medical students do not understand how to use email for personal communication due to unfamiliarity; instead they try to use it as if it were Facebook or Twitter.

In addition, professionalism is acquired over time and is best learned within the practice community and specifically through observation of role models [46]. However, mentoring and observation of role models as a vital component of developing professionalism might face difficulties in the digital age, with different generations of physicians practicing in parallel [23]. This particular situation further favors policies that capture the broad spectrum of challenges and opportunities for medical professionalism with respect to social media.

Limitations

There are some limitations to our review: we screened only contributions published (in different types of publications) in scientific journals listed in PubMed. Only German and English publications were considered. Only publications after 2002 were included, due to the fact that all major social media platforms

were founded after the year 2002 [18-20]. While our search revealed 1471 references listed in PubMed for the years 2002-2011, another 982 references are listed in PubMed in 2012 that could not be included in this review. Because our review already included more than 100 references published in journals from various subspecialties and because we reached theoretical saturation for our broad categories of opportunities and challenges, we felt justified in limiting our review to the described literature search.

Because the findings of our review are purely descriptive and we did not provide additional normative analysis to each of the identified challenges and opportunities, we refrain from concluding on how these challenges and opportunities should be best addressed in medical practice. However, the recently published position paper by the American College of Physicians and the Federation of State Medical Board presents several distinguished implications of online activities for patients, physicians, and the medical profession and provides recommendations on how to avoid potential pitfalls while best using social media technologies [17]. Also, other in-depth analyses result in specific suggestions on how to deal with social media-related challenges and opportunities [16,47]. However, none of the above mentioned policy and recommendation papers refer to a systematically and transparently derived account of challenges and opportunities.

Conclusions

The integration of traditional core values of medicine (privacy, confidentiality, one-on-one interactions, and formal conduct) and the culture of social media (which tends to value sharing and openness, connection, transparency, and informality) present opportunities as well as challenges for medical professionalism [24]. As a profession that is entitled to self-regulation, health care professionals should proactively approach these challenges and make use of the opportunities. There should be room for fostering interprofessional and intergenerational dialogue (eg, digital natives/digital immigrants). There is a further demand for research and policy development to integrate the broad spectrum of social media's opportunities and challenges into the current existing frameworks for medical professionalism. This review builds a unique source of information that can inform further research and policy development in this regard.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The spectrum of social media-related opportunities and challenges for medical professionalism.

[\[PDF File \(Adobe PDF File\), 134KB - jmir_v15i8e184_app1.pdf\]](#)

References

1. ABIM Foundation. American Board of Internal Medicine, ACP-ASIM Foundation. American College of Physicians-American Society of Internal Medicine, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med* 2002;136(3):243-246. [Medline: [11827500](#)]

2. Lau AS. Hospital-based nurses' perceptions of the adoption of Web 2.0 tools for knowledge sharing, learning, social interaction and the production of collective intelligence. *J Med Internet Res* 2011;13(4):e92 [FREE Full text] [doi: [10.2196/jmir.1398](https://doi.org/10.2196/jmir.1398)] [Medline: [22079851](https://pubmed.ncbi.nlm.nih.gov/22079851/)]
3. Thompson LA, Dawson K, Ferdig R, Black EW, Boyer J, Coutts J, et al. The intersection of online social networking with medical professionalism. *J Gen Intern Med* 2008 Jul;23(7):954-957 [FREE Full text] [doi: [10.1007/s11606-008-0538-8](https://doi.org/10.1007/s11606-008-0538-8)] [Medline: [18612723](https://pubmed.ncbi.nlm.nih.gov/18612723/)]
4. Gabbard GO, Kassaw KA, Perez-Garcia G. Professional boundaries in the era of the Internet. *Acad Psychiatry* 2011;35(3):168-174. [doi: [10.1176/appi.ap.35.3.168](https://doi.org/10.1176/appi.ap.35.3.168)] [Medline: [21602438](https://pubmed.ncbi.nlm.nih.gov/21602438/)]
5. Freebury R, Ennis J, Rideout C, Wright M. General Guidelines for the Practice of Psychotherapy. In: Cameron P, Ennis J, Deadman J, editors. *Standards and Guidelines for the Psychotherapies*. Toronto: University of Toronto Press; 1998:17-43.
6. Thompson LA, Black E, Duff WP, Paradise Black N, Saliba H, Dawson K. Protected health information on social networking sites: ethical and legal considerations. *J Med Internet Res* 2011 Jan;13(1):e8 [FREE Full text] [doi: [10.2196/jmir.1590](https://doi.org/10.2196/jmir.1590)] [Medline: [21247862](https://pubmed.ncbi.nlm.nih.gov/21247862/)]
7. Boyd D. Social Network Sites: Public, Private or What?. 2012. URL: <http://kt.flexiblelearning.net.au/tkt2007/edition-13/> [accessed 2013-04-07] [WebCite Cache ID 6FhxdA3F1]
8. Mansfield SJ, Morrison SG, Stephens HO, Bonning MA, Wang SH, Withers AH, et al. Social media and the medical profession. *Med J Aust* 2011 Jun 20;194(12):642-644. [Medline: [21692723](https://pubmed.ncbi.nlm.nih.gov/21692723/)]
9. Chretien KC, Greysen SR, Chretien JP, Kind T. Online posting of unprofessional content by medical students. *JAMA* 2009 Sep 23;302(12):1309-1315. [doi: [10.1001/jama.2009.1387](https://doi.org/10.1001/jama.2009.1387)] [Medline: [19773566](https://pubmed.ncbi.nlm.nih.gov/19773566/)]
10. Griffith R, Tegnah C. District nurses' use of social networking sites: caution required. *Br J Community Nurs* 2011 Sep;16(9):455-457. [Medline: [22067956](https://pubmed.ncbi.nlm.nih.gov/22067956/)]
11. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. *J Med Internet Res* 2010;12(2):e19 [FREE Full text] [doi: [10.2196/jmir.1549](https://doi.org/10.2196/jmir.1549)] [Medline: [20542858](https://pubmed.ncbi.nlm.nih.gov/20542858/)]
12. Lagu T, Kaufman EJ, Asch DA, Armstrong K. Content of weblogs written by health professionals. *J Gen Intern Med* 2008 Oct;23(10):1642-1646 [FREE Full text] [doi: [10.1007/s11606-008-0726-6](https://doi.org/10.1007/s11606-008-0726-6)] [Medline: [18649110](https://pubmed.ncbi.nlm.nih.gov/18649110/)]
13. Lee K, Seda C. *Search Engine Advertising: Buying your way to the top to increase sales*. Berkeley, CA: New Riders; 2009.
14. Burns A, Wolstencroft S. The internet: education, social media and dental practice. *J Ir Dent Assoc* 2011;57(5):268-271. [Medline: [22165478](https://pubmed.ncbi.nlm.nih.gov/22165478/)]
15. Syed-Abdul S, Fernandez-Luque L, Jian WS, Li YC, Crain S, Hsu MH, et al. Misleading health-related information promoted through video-based social media: anorexia on YouTube. *J Med Internet Res* 2013;15(2):e30 [FREE Full text] [doi: [10.2196/jmir.2237](https://doi.org/10.2196/jmir.2237)] [Medline: [23406655](https://pubmed.ncbi.nlm.nih.gov/23406655/)]
16. Chretien KC, Kind T. Social media and clinical care: ethical, professional, and social implications. *Circulation* 2013 Apr 2;127(13):1413-1421. [doi: [10.1161/CIRCULATIONAHA.112.128017](https://doi.org/10.1161/CIRCULATIONAHA.112.128017)] [Medline: [23547180](https://pubmed.ncbi.nlm.nih.gov/23547180/)]
17. Farnan JM, Snyder Sulmasy L, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM, American College of Physicians Ethics, Professionalism Human Rights Committee, American College of Physicians Council of Associates, Federation of State Medical Boards Special Committee on Ethics Professionalism*. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. *Ann Intern Med* 2013 Apr 16;158(8):620-627. [doi: [10.7326/0003-4819-158-8-201304160-00100](https://doi.org/10.7326/0003-4819-158-8-201304160-00100)] [Medline: [23579867](https://pubmed.ncbi.nlm.nih.gov/23579867/)]
18. Wikipedia. Myspace. URL: <http://de.wikipedia.org/wiki/Myspace> [accessed 2013-04-07] [WebCite Cache ID 6Fhy4WIHt]
19. Facebook. Facebook News Room. URL: <http://newsroom.fb.com/content/default.aspx?NewsAreaId=22> [accessed 2013-04-07] [WebCite Cache ID 6Fi0QiqFg]
20. Wikipedia. Twitter. URL: <http://de.wikipedia.org/wiki/Twitter> [accessed 2013-04-07] [WebCite Cache ID 6Fi0j8AC8]
21. Grodin M. Introduction: the historical and philosophical roots of bioethics. In: *Meta medical ethics: the philosophical foundations of bioethics*. London: Kluwer Academic Publishers; 1995.
22. Hyman JL, Luks HJ, Sechrest R. Online professional networks for physicians: risk management. *Clin Orthop Relat Res* 2012 May;470(5):1386-1392. [doi: [10.1007/s11999-011-2197-z](https://doi.org/10.1007/s11999-011-2197-z)] [Medline: [22125249](https://pubmed.ncbi.nlm.nih.gov/22125249/)]
23. Gorrindo T, Groves JE. Medical professionalism: a tale of two doctors. *J Clin Ethics* 2011;22(2):176-178. [Medline: [21837890](https://pubmed.ncbi.nlm.nih.gov/21837890/)]
24. George DR. "Friending Facebook?" A minicourse on the use of social media by health professionals. *J Contin Educ Health Prof* 2011;31(3):215-219. [doi: [10.1002/chp.20129](https://doi.org/10.1002/chp.20129)] [Medline: [21953663](https://pubmed.ncbi.nlm.nih.gov/21953663/)]
25. Farnan JM, Arora VM. Blurring boundaries and online opportunities. *J Clin Ethics* 2011;22(2):183-186. [Medline: [21837892](https://pubmed.ncbi.nlm.nih.gov/21837892/)]
26. Swartz MK. The potential for social media. *J Pediatr Health Care* 2011;25(6):345. [doi: [10.1016/j.pedhc.2011.08.005](https://doi.org/10.1016/j.pedhc.2011.08.005)] [Medline: [22018424](https://pubmed.ncbi.nlm.nih.gov/22018424/)]
27. Tilt A, Mermel C, Conrad C. How surgical residents use social media. *Surgery* 2011 Jul;150(1):5-6. [doi: [10.1016/j.surg.2011.05.022](https://doi.org/10.1016/j.surg.2011.05.022)] [Medline: [21683854](https://pubmed.ncbi.nlm.nih.gov/21683854/)]
28. Kukreja P, Heck Sheehan A, Riggins J. Use of social media by pharmacy preceptors. *Am J Pharm Educ* 2011 Nov 10;75(9):176 [FREE Full text] [doi: [10.5688/ajpe759176](https://doi.org/10.5688/ajpe759176)] [Medline: [22171104](https://pubmed.ncbi.nlm.nih.gov/22171104/)]
29. Varga-Atkins T, Dangerfield P, Brigden D. Developing professionalism through the use of wikis: A study with first-year undergraduate medical students. *Med Teach* 2010;32(10):824-829. [doi: [10.3109/01421591003686245](https://doi.org/10.3109/01421591003686245)] [Medline: [20854158](https://pubmed.ncbi.nlm.nih.gov/20854158/)]

30. Coe JB, Weijs CA, Muise A, Christofides E, Desmarais S. Teaching veterinary professionalism in the Face(book) of change. *J Vet Med Educ* 2011;38(4):353-359. [doi: [10.3138/jvme.38.4.353](https://doi.org/10.3138/jvme.38.4.353)] [Medline: [22130411](https://pubmed.ncbi.nlm.nih.gov/22130411/)]
31. Jent JF, Eaton CK, Merrick MT, Englebort NE, Dandes SK, Chapman AV, et al. The decision to access patient information from a social media site: what would you do? *J Adolesc Health* 2011 Oct;49(4):414-420 [FREE Full text] [doi: [10.1016/j.jadohealth.2011.02.004](https://doi.org/10.1016/j.jadohealth.2011.02.004)] [Medline: [21939873](https://pubmed.ncbi.nlm.nih.gov/21939873/)]
32. Bongartz J, Vang C, Havrda D, Fravel M, McDanel D, Farris KB. Student pharmacist, pharmacy resident, and graduate student perceptions of social interactions with faculty members. *Am J Pharm Educ* 2011 Nov 10;75(9):180 [FREE Full text] [doi: [10.5688/ajpe759180](https://doi.org/10.5688/ajpe759180)] [Medline: [22171108](https://pubmed.ncbi.nlm.nih.gov/22171108/)]
33. Go PH, Klaassen Z, Chamberlain RS. Residency selection: do the perceptions of US programme directors and applicants match? *Med Educ* 2012 May;46(5):491-500. [doi: [10.1111/j.1365-2923.2012.04257.x](https://doi.org/10.1111/j.1365-2923.2012.04257.x)] [Medline: [22515757](https://pubmed.ncbi.nlm.nih.gov/22515757/)]
34. Shore R, Halsey J, Shah K, Crigger BJ, Douglas SP, AMA Council on Ethical and Judicial Affairs (CEJA). Report of the AMA Council on Ethical and Judicial Affairs: professionalism in the use of social media. *J Clin Ethics* 2011;22(2):165-172. [Medline: [21837888](https://pubmed.ncbi.nlm.nih.gov/21837888/)]
35. McCartney M. Online is not a private space. *BMJ* 2011 Nov 02;343:d7109. [doi: [10.1136/bmj.d7109](https://doi.org/10.1136/bmj.d7109)]
36. Thielst CB. Using social media to engage patients: many tools exist to connect, communicate and build loyalty. *Healthc Exec* 2011;26(3):66, 68-66, 70. [Medline: [21612166](https://pubmed.ncbi.nlm.nih.gov/21612166/)]
37. Prasanna PM, Seagull FJ, Nagy P. Online social networking: a primer for radiology. *J Digit Imaging* 2011 Oct;24(5):908-912 [FREE Full text] [doi: [10.1007/s10278-011-9371-4](https://doi.org/10.1007/s10278-011-9371-4)] [Medline: [21360214](https://pubmed.ncbi.nlm.nih.gov/21360214/)]
38. Omurtag K, Jimenez PT, Ratts V, Odem R, Cooper AR. The ART of social networking: how SART member clinics are connecting with patients online. *Fertil Steril* 2012 Jan;97(1):88-94. [doi: [10.1016/j.fertnstert.2011.10.001](https://doi.org/10.1016/j.fertnstert.2011.10.001)] [Medline: [22088209](https://pubmed.ncbi.nlm.nih.gov/22088209/)]
39. Knösel M, Jung K, Bleckmann A. YouTube, dentistry, and dental education. *J Dent Educ* 2011 Dec;75(12):1558-1568 [FREE Full text] [Medline: [22184594](https://pubmed.ncbi.nlm.nih.gov/22184594/)]
40. Williams J, Feild C, James K. The effects of a social media policy on pharmacy students' facebook security settings. *Am J Pharm Educ* 2011 Nov 10;75(9):177 [FREE Full text] [doi: [10.5688/ajpe759177](https://doi.org/10.5688/ajpe759177)] [Medline: [22171105](https://pubmed.ncbi.nlm.nih.gov/22171105/)]
41. 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/)]
42. Kind T, Genrich G, Sodhi A, Chretien KC. Social media policies at US medical schools. *Med Educ Online* 2010;15:- [FREE Full text] [doi: [10.3402/meo.v15i0.5324](https://doi.org/10.3402/meo.v15i0.5324)] [Medline: [20859533](https://pubmed.ncbi.nlm.nih.gov/20859533/)]
43. American Medical Association. Opinion 9.124 - Professionalism in the Use of Social Media. URL: <http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion9124.page> [accessed 2013-04-07] [WebCite Cache ID 6Fi1D84iR]
44. UF Health. Use of Social Networking Sites. Florida: Uo URL: <http://osa.med.ufl.edu/policies/use-of-social-networking-sites/> [accessed 2013-04-07] [WebCite Cache ID 6Fi1rfrce]
45. Prensky M. Digital Natives, Digital Immigrants. In: *On the Horizon*, vol 9 no 5. Lincoln: NCB University Press; 2001:1-6.
46. Finn G, Garner J, Sawdon M. 'You're judged all the time!' Students' views on professionalism: a multicentre study. *Med Educ* 2010 Aug;44(8):814-825. [doi: [10.1111/j.1365-2923.2010.03743.x](https://doi.org/10.1111/j.1365-2923.2010.03743.x)] [Medline: [20633221](https://pubmed.ncbi.nlm.nih.gov/20633221/)]
47. Greysen SR, Kind T, Chretien KC. Online professionalism and the mirror of social media. *J Gen Intern Med* 2010 Nov;25(11):1227-1229 [FREE Full text] [doi: [10.1007/s11606-010-1447-1](https://doi.org/10.1007/s11606-010-1447-1)] [Medline: [20632121](https://pubmed.ncbi.nlm.nih.gov/20632121/)]

Abbreviations

ABIM: American Board of Internal Medicine

ACP-ASIM: American College of Physicians-American Society of Internal Medicine

AMA: American Medical Association

Edited by G Eysenbach; submitted 16.05.13; peer-reviewed by J Fernan, M Broom; comments to author 08.06.13; revised version received 20.06.13; accepted 20.06.13; published 28.08.13.

Please cite as:

Gholami-Kordkheili F, Wild V, Strech D

The Impact of Social Media on Medical Professionalism: A Systematic Qualitative Review of Challenges and Opportunities

J Med Internet Res 2013;15(8):e184

URL: <http://www.jmir.org/2013/8/e184/>

doi: [10.2196/jmir.2708](https://doi.org/10.2196/jmir.2708)

PMID: [23985172](https://pubmed.ncbi.nlm.nih.gov/23985172/)

©Fatemeh Gholami-Kordkheili, Verina Wild, Daniel Strech. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.08.2013. 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

The Clinical Effectiveness of Web-Based Cognitive Behavioral Therapy With Face-to-Face Therapist Support for Depressed Primary Care Patients: Randomized Controlled Trial

Ragnhild Sørensen Høifødt¹, PsyD; Kjersti R Lillevoll¹, PsyD; Kathleen M Griffiths², PhD; Tom Wilsgaard³, PhD; Martin Eisemann¹, PhD; Knut Waterloo¹, PhD; Nils Kolstrup⁴, MD, PhD

¹Department of Psychology, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway

²Centre for Mental Health Research, The Australian National University, Canberra, Australia

³Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway

⁴Department of Community Medicine, General Practice Research Unit, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway

Corresponding Author:

Ragnhild Sørensen Høifødt, PsyD

Department of Psychology

Faculty of Health Sciences

University of Tromsø

Tromsø, 9037

Norway

Phone: 47 776 49230

Fax: 47 776 45291

Email: ragnhild.s.hoifodt@uit.no

Abstract

Background: Most patients with mild to moderate depression receive treatment in primary care, but despite guideline recommendations, structured psychological interventions are infrequently delivered. Research supports the effectiveness of Internet-based treatment for depression; however, few trials have studied the effect of the MoodGYM program plus therapist support. The use of such interventions could improve the delivery of treatment in primary care.

Objective: To evaluate the effectiveness and acceptability of a guided Web-based intervention for mild to moderate depression, which could be suitable for implementation in general practice.

Methods: Participants (N=106) aged between 18 and 65 years were recruited from primary care and randomly allocated to a treatment condition comprising 6 weeks of therapist-assisted Web-based cognitive behavioral therapy (CBT), or to a 6-week delayed treatment condition. The intervention included the Norwegian version of the MoodGYM program, brief face-to-face support from a psychologist, and reminder emails. The primary outcome measure, depression symptoms, was measured by the Beck Depression Inventory-II (BDI-II). Secondary outcome measures included the Beck Anxiety Inventory (BAI), the Hospital Anxiety and Depression Scale (HADS), the Satisfaction with Life Scale (SWLS), and the EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D). All outcomes were based on self-report and were assessed at baseline, postintervention, and at 6-month follow-up.

Results: Postintervention measures were completed by 37 (71%) and 47 (87%) of the 52 participants in the intervention and 54 participants in the delayed treatment group, respectively. Linear mixed-models analyses revealed a significant difference in time trends between the groups for the BDI-II, ($P=.002$), for HADS depression and anxiety subscales ($P<.001$ and $P=.001$, respectively), and for the SWLS ($P<.001$). No differential group effects were found for the BAI and the EQ-5D. In comparison to the control group, significantly more participants in the intervention group experienced recovery from depression as measured by the BDI-II. Of the 52 participants in the treatment program, 31 (60%) adhered to the program, and overall treatment satisfaction was high. The reduction of depression and anxiety symptoms was largely maintained at 6-month follow-up, and positive gains in life satisfaction were partly maintained.

Conclusions: The intervention combining MoodGYM and brief therapist support can be an effective treatment of depression in a sample of primary care patients. The intervention alleviates depressive symptoms and has a significant positive effect on anxiety symptoms and satisfaction with life. Moderate rates of nonadherence and predominately positive evaluations of the

treatment also indicate the acceptability of the intervention. The intervention could potentially be used in a stepped-care approach, but remains to be tested in regular primary health care.

Trial Registration: Australian New Zealand Clinical Trials Registry: ACTRN12610000257066; <http://apps.who.int/trialsearch/trial.aspx?trialid=ACTRN12610000257066> (Archived by WebCite at <http://www.webcitation.org/6Ie3YhIZa>).

(*J Med Internet Res* 2013;15(8):e153) doi:[10.2196/jmir.2714](https://doi.org/10.2196/jmir.2714)

KEYWORDS

cognitive therapy; therapy; computer-assisted; Internet; mental health; depression; randomized controlled trial; primary health care

Introduction

Overview

Depression is a highly prevalent disorder that often causes substantial functional impairment in daily life, reduction in quality of life, and increased medical service utilization [1-6]. There exist several effective psychological and pharmacological treatments for depression [7]. However, a large proportion of those suffering from this disorder receive inadequate treatment or no treatment at all [8,9]. Cognitive behavioral therapy (CBT) has proven to be as effective as pharmacotherapy in treating mild to moderate depression, with the benefit of reduced rates of relapse [10,11].

Internet-Based Treatment of Depression

The principles and techniques of CBT have been extensively disseminated through self-help books and computer- or Internet-based programs. A substantive body of research shows that Internet-based CBT can be an efficacious treatment of depression (eg, [12-15]). Research also suggests that such interventions are cost-effective compared to face-to-face treatments because they result in symptom reduction and reduced burden of disease for patients and alleviate demands on clinician time and resources [16-18].

Self-help can be self-administered or guided by a therapist, although the active involvement of the therapist in guided self-help is less extensive than in conventional psychological therapy. Studies generally show small to moderate effects of self-administered, unguided CBT in the treatment of depression [19-23], although in some studies unguided interventions have yielded large treatment effects [24,25]. Still, an increasing amount of research has pointed to the importance of support in Internet-based interventions, with interventions offering some degree of support from a professional during treatment generally showing substantially larger treatment effects than interventions involving little or no professional support [26-28]. However, this conclusion is primarily based on meta-analytic results; the results from the few studies directly comparing guided and unguided interventions are mixed [14,24,29]. Overall, guided interventions show moderate to large treatment effects for depression, and the average effect sizes for guided self-help are comparable to the effects of time-limited face-to-face treatment (eg, [13-15,30]). This is further supported by a recent meta-analysis, which found no significant differences in effect between guided self-help and face-to-face therapy [31].

MoodGYM is a free Web-based CBT program developed to prevent and treat mild to moderate depression [32]. Studies have demonstrated the effectiveness of MoodGYM in reducing symptoms of depression and anxiety among public registrants, trial participants, callers to a national helpline service, users of the UK National Health Service portal, adolescent school-based populations, and in Australian and Norwegian student samples [20,24,25,29,33-35]. Positive effects have been shown to be sustained over 12 months [36]. However, few previous trials have investigated the effect of MoodGYM combined with therapist support. A study found that the conjunction of MoodGYM and face-to-face therapy was superior to both MoodGYM alone and for some outcome measures, to time-limited face-to-face therapy alone [29]. Also the results of a cluster-randomized trial suggested positive effects of the combination of MoodGYM and support from general practitioners (GPs) compared to GP care alone [37].

Depression Treatment in Primary Health Care

Most patients with psychological problems will receive most or all of their mental health care in primary care, and findings suggest that many patients prefer to consult their GP for treatment of depression [2,38-40]. Clinical practice guidelines primarily recommend treating mild to moderate depression using psychosocial interventions [41,42], and this is also in accordance with reported patient preferences [43-45]. Nevertheless, structured psychological interventions are infrequently delivered in general practice [46-48] because of time constraints [49-51] and a lack of knowledge and competence among GPs in the delivery of evidence-based psychological interventions [51,52]. The use of CBT-based self-help resources could be a way to improve the delivery of psychological interventions in general practice. This would allow for short consultations and for the clinician to be a facilitator rather than a cognitive therapist. These features could improve feasibility in general practice, where the volume of patients is high and it is essential that interventions are brief and practical.

Aim of the Study

The current study was designed to trial a procedure for depression treatment that could be suitable for implementation in general practice. The project was planned as the first phase of research for this treatment, with the second phase focusing on further evaluation carried out in everyday general practice. The aim was to evaluate the effectiveness and acceptability of a guided self-help intervention combining the MoodGYM

program with brief face-to-face therapist support in a sample of primary care patients with mild to moderate symptoms of depression. This was investigated in a randomized controlled trial comparing the guided self-help intervention to a delayed-treatment control condition. The primary hypothesis was that therapist-supported Web-based CBT would lead to a larger reduction in depressive symptoms than the control condition. To determine if the intervention was acceptable to patients, satisfaction with treatment, adherence, and reasons for dropout were investigated.

Methods

Study Design

The study was a randomized controlled trial with balanced randomization (1:1). Participants were randomly allocated to a treatment condition comprising 6 weeks of Web-based CBT with therapist support, or to a 6-week waitlist for the same treatment during which time they could also access treatment as usual. The study was conducted at the Department of Psychology at the University of Tromsø where a small self-help outpatient clinic was established. The research protocol was approved by the Regional Committee for Research Ethics in Northern Norway (2011/2163) and the Human Ethics Committee of the Australian National University (ANU). The trial was registered in the Australian New Zealand Clinical Trials Registry (ACTRN1261000257066). The trial is reported in accordance with the CONSORT-EHEALTH [53] (see [Multimedia Appendices 1-3](#)).

Participants

Participants (N=106) were recruited between October 2010 and October 2012 from GPs, primary care nurses, and from waitlists of primary care referrals at 2 psychiatric outpatient clinics. Calculations of required sample size were based on a power of .80, significance level of .05 (2-sided), and an expected effect size of 0.6 on depressive symptoms at posttest. The estimations necessitated a sample size of 45 participants per group. A median dropout rate between 17% and 19% has been reported for computerized or Web-based treatment programs [54,55]. With a 20% expected dropout, a total sample size of 108 was required to gain sufficient power, yielding group sizes of 54 participants.

Local GPs and primary care nurses were informed about the study both verbally at practice meetings and through written information. They provided patients who they considered as mildly to moderately depressed based on clinical appraisal and/or screening instruments with written information about the project. Potentially eligible patients on waitlists for psychiatric outpatient treatment were identified by clinic staff and subsequently received information by postal mail from the research group. When informing a patient of the project, all recruiters were asked to send a notification to the research group by using a prepaid envelope. The notification simply notified the researchers that a patient had been informed of the project and did not reveal any information about the patient. Patients were provided with general information about the treatment and the aim of the project and detailed information about the methods for handling issues of privacy and anonymity. They

were informed that they could expect to commence treatment within 6 weeks of the initial contact. To participate, patients sent in a signed informed consent form providing contact details. Study inclusion criteria were: (1) age 18-65 years, (2) access to the Internet, and (3) a score between 14 and 29 on the Beck Depression Inventory-II (BDI-II), indicating mild to moderate symptoms of depression. During the first months of the study, the protocol was changed by extending the inclusion criterion on the BDI-II to include participants with scores between 10 and 40. This change was because of insufficient recruitment and the clinical appraisal that patients with scores above 30 could possibly benefit from the treatment based on their daily functioning and motivation. In addition, their depression was too mild to assure them other public treatment options. Furthermore, several patients with a BDI-II score below 14 reported a need for treatment. Based on this revised criteria, 7 eligible patients were falsely excluded in the initial phase of the trial. Individuals currently undergoing CBT were excluded, whereas individuals who used antidepressant medication were stabilized for 1 month prior to evaluation of diagnostic eligibility. To maximize the external validity of the trial, a heterogeneous group of patients with depressive symptoms was included, independent of a particular diagnosis. Therefore, medical or psychiatric comorbidities only restricted inclusion when there was a need for immediate treatment of these comorbid conditions (suicidal ideation, current psychosis) or if the conditions were expected to markedly interfere with treatment of the depressive condition (alcohol or drug use disorders).

Measures

All outcome measures were based on self-report. Assessments of symptoms of depression, anxiety, and quality of life using paper-and-pencil questionnaires were completed by all participants at baseline, posttreatment, and at 6 months posttreatment (online questionnaires). The control group also completed these inventories before entering online treatment (postwaiting). BDI-II was administered before every consultation during the intervention phase.

The primary outcome measure was the BDI-II, a 21-item measure of severity of depressive symptoms during the past 2 weeks [56]. Each item is rated on a 4-point scale ranging from 0 to 3. Studies consistently support the BDI-II as a reliable, internally consistent, and valid scale for assessing depression both in psychiatric outpatients, the general population, and in primary care settings [56-58]. Several studies have found high correlations between the paper-and-pencil and the computerized/online versions of the BDI-II [59-62]. In the present study, internal consistency (Cronbach alpha) was .78 pretreatment, .93 posttreatment, and .94 at 6-month follow-up, respectively.

The secondary outcome measures consisted of the Beck Anxiety Inventory (BAI), the Hospital Anxiety and Depression Scale (HADS), and 2 measures of quality of life—Satisfaction With Life Scale (SWLS) and the EuroQol Group 5-Dimension Self-Report Questionnaire (EQ-5D)—as well as a measure of treatment satisfaction. The quality of life measures were included during the initial phase of the trial considering that the

extension of outcomes beyond symptom measures would strengthen the study.

The BAI is a 21-item measure of anxiety symptom severity [63]. Each item is rated from 0 to 3 depending on symptom severity during the past week. The inventory possesses high internal consistency and reliability, as well as robust convergent and discriminant validity [64-66]. Equivalent psychometric properties have been shown across paper-and-pencil and online formats of the questionnaire, and the 2 formats are highly correlated [59,67]. Cronbach alphas in the present study were .93 at pretreatment, .88 at posttreatment, and .92 at 6-month follow-up, respectively.

The HADS is a 14-item inventory with 2 subscales of 7 items each, measuring depression and anxiety, respectively [68]. Each item is rated on a 0 to 3 scale. The inventory has good to very good construct validity and internal consistency [68-70]. Most factor analyses confirm a 2-factor solution comprising a depression and an anxiety subscale [69]. Paper-and-pencil and Internet administration of the measure yields comparable psychometric properties, but Internet administration may overestimate scores [70,71]. In the present study, Cronbach alpha was .68 and .82 at pretreatment, .82 and .84 at posttreatment, and .87 and .86 at 6-month follow-up for the depression and anxiety subscales, respectively.

The EQ-5D is a generic questionnaire evaluating health-related quality of life [72]. The respondent marks his/her level of functioning (no problems, some problems, extreme problems) for each of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and rates his/her health state on a visual analog scale (EQ VAS) with the endpoints labeled best imaginable health state and worst imaginable health state, respectively. For the 5 dimensions, a scoring algorithm (the MVH_A1 tariff) based on preference weights was used to aggregate an index score (EQ Index) [73]. The health states "perfect health" (no problems on any dimension) and "dead" are assigned the values of 1 and 0, respectively. The instrument has been demonstrated to discriminate between subgroups of patients with differing severities of mental illness and to capture changes in quality of life associated with improved mental health over time [74]. No significant differences have been found between scores obtained using paper and computerized modes of administration [75,76].

The SWLS measures global life satisfaction as a cognitive-judgemental process, in which individuals assess their quality of life according to their own criteria [77]. The respondent rates on a 7-point Likert scale the degree to which he/she agrees with 5 statements. Several studies confirm the scale's unidimensional structure and support its sound psychometric properties, including internal consistency, test-retest reliability, as well as convergent and discriminant construct validity [78-80]. Research indicates that Internet data on this measure is as reliable and valid as paper-and-pencil data [81]. Cronbach alpha in this study was .79 at pretreatment, .87 at posttreatment, and .93 at 6-month follow-up.

The Mini-International Neuropsychiatric Interview (MINI) is a short, structured diagnostic interview for identifying the diagnoses of the *Diagnostic and Statistical Manual of Mental*

Disorders (Fourth Edition; *DSM-IV*) and the *International Classification of Diseases, Tenth Revision (ICD-10)*. It consists of 17 modules. A comparison of MINI with other structured clinical interviews shows sensitivities and specificities above 0.70 for most diagnoses [82]. Excellent interrater reliability has been reported. The MINI Interview was used to determine psychiatric comorbidity and for excluding participants with current psychosis or suicidal ideation, as indicated by 17 points or more on the suicidal ideation module.

The Alcohol Use Disorders Identification Test (AUDIT) is a screening instrument consisting of 10 questions about alcohol use in the past 12 months, alcohol dependence symptoms, and alcohol-related problems [83]. Eight items are rated on a 5-point scale (0-4) and 2 items are rated on a 3-point scale (0, 2, 4). A large body of research confirms the favorable internal consistency, reliability, and criterion validity [83-85]. In this study, the scale was used to screen for alcohol use problems. A cutoff score of 20 was chosen to exclude patients in need of further diagnostic evaluation for alcohol dependence [85]. For participants scoring above 16, alcohol use was monitored during treatment.

The Drug Use Disorders Identification Test (DUDIT) is an 11-item screening instrument measuring patterns of substance use during the past 12 months, as well as various drug-related problems [86]. Nine items are rated on a 5-point scale (0-4) and 2 items are rated on a 3-point scale (0, 2, 4). In a sample of drug users, the scale has shown good reliability, and it predicts drug dependence with a sensitivity of 0.90 and average specificity of approximately 0.80 [86]. In the present study, the DUDIT was used to screen for drug use disorder. A cutoff score of 25 was used to exclude patients with a high probability of drug dependency.

Satisfaction with treatment was measured by 9 questions that respondents rated on a 5-point scale (1-5, very negative to very positive). The questions concerned their satisfaction with the intervention as a whole and various aspects of the self-help program and follow-up sessions. The general content of the questions was influenced by patient satisfaction questionnaires used in other studies [87-89]. However, the exact content was tailored to tap into aspects of treatment considered important for the purpose of the present investigation. The questions are described in detail in the Results section.

Treatment variables included module completion, number of sessions, treatment duration, session duration (in minutes, not including screening), and total time spent by therapists (time spent outside the consultations was not registered). User data on module completion was registered online and was denoted by a number between 0 and 4, with 0 indicating no use and 4 indicating completion of the module. For the variable time spent by therapists, the amount of missing data was considerable (51.9% for time spent on screening, 14.6% for total time); thus, these data can only be considered estimates. Total time was estimated by imputing mean screening time for missing data concerning screening duration and each individual's mean session duration for missing data from treatment sessions.

Procedure

After informed consent was obtained, participants were screened for inclusion through a face-to-face session. A computerized random number generator randomized identification (ID) numbers to the 2 groups (generated by KL). Eligible participants were given ID numbers following a chronological sequence. To ensure equal group sizes, blocked randomization with variable block sizes was used. Patients could not be blinded to group assignment, but were blinded to the status of the waitlist as a control condition.

Screening, enrollment, and treatment were carried out by 2 licensed clinical psychologists (RSH and KL) with basic CBT skills and good knowledge of the MoodGYM program. Both had less than 2 years of experience in clinical practice and no prior experience with Internet-based treatment. The therapists were not blind to the participants' group allocation. However, steps were taken to blind the evaluation of outcomes by ensuring that posttests were performed by a research assistant unaware of the participants' allocation assignment.

Intervention

Participants in both groups were free to access usual primary care treatment, which could include antidepressant medication, informal supportive therapy, or referral to specialist mental health services.

The guided self-help intervention involved 3 components: (1) The Norwegian version of the Web-based CBT program MoodGYM version 3 [90], (2) brief face-to-face therapist support, and (3) tailored emails between sessions. MoodGYM was originally developed at the Australian National University to prevent depression in young people aged between 15 and 25 years. However, data from individuals who used the English version of the program has shown that most users were aged 25 to 44 years, and that the users' average depression and anxiety scores were elevated compared to the general population [91]. Therefore, the program appears useful for older age groups than originally targeted, and for individuals with elevated levels of depressive and anxious symptomatology. The program consists of 5 modules and a personal workbook containing exercises and assessments. Module 1 through 3 focus on the cognitive model, typical patterns of dysfunctional thinking, and exercises to identify and restructure dysfunctional thinking, as well as behavioral strategies to increase engagement in positive activities. Module 4 focuses on stress and stress reduction and introduces relaxation techniques. Module 5 covers simple problem solving and typical responses to broken relationships. Each module takes approximately 45 to 60 minutes to work through. See [Figure 1](#) for screenshots from the program.

In the first session after screening, participants were introduced to the program, received their trial username and password, and were instructed to work at home with 1 module each week. After each module, participants received face-to-face support (15-30 minutes). The therapists followed a guideline script with 3 compulsory topics for every consultation: (1) monitoring of depression symptoms and discussion of changes, (2) a focus on the important topics and exercises covered by each module and the participants' experiences of working with it, and (3)

introduction of the next module and motivate participants to adhere to the treatment plan. The main focus of the therapist was on reinforcing the efforts made by participants and helping them to relate to the material and to incorporate the use of techniques from the program into their everyday living. If time permitted, participants could also bring up other topics they considered important in relation to their depression. In the concluding session, the experiences and outcomes of treatment were discussed. Therapists aimed to meet participants weekly and to complete the intervention over 7 weeks. However, the interval between sessions and the number of sessions were allowed to vary somewhat to provide flexibility in meeting individual needs. Between sessions, participants received tailored emails aiming to motivate them to work with the self-help program. The emails introduced the next module, and some contained brief advice on how to overcome depressive symptoms (eg, the importance of behavioral activation). Participants did not get a mental health record at the clinic, but a short case summary was sent to their GP (with consent from the participants). Participants considered in need of more extensive treatment throughout or after completing the trial were assisted in the process of referral to specialized mental health services. The Web-based program did not store any personally identifying information about users.

Statistical Analyses

All analyses were carried out using IBM SPSS Statistics version 19 for Windows (IBM Corp, Armonk, NY, USA), except for the power calculation which was performed using G*Power (Heinrich-Heine-Universität, Institut für Experimentelle Psychologie, Düsseldorf, Germany).

Differences between the groups on baseline characteristics were examined by performing chi-square tests for categorical variables and 1-way analysis of variance (ANOVA) for continuous variables. A logistic regression analysis with backward stepwise method was used to explore whether missing data at postintervention (not completing postintervention measures) could be significantly predicted by participant characteristics.

Results on the BDI-II and BAI were analyzed using intention-to-treat (ITT) analyses in which participants are analyzed in the group they were randomized to, irrespective of treatment adherence. Because of missing data at pretest for the remaining secondary measures (3% missing on the HADS, and 17% and 19% missing on SWLS and EQ-5D, respectively), modified ITT analysis was performed including all participants completing the measures at least once. Effects were tested by performing linear mixed-models analysis using the restricted maximum likelihood (REML) estimation procedure and an unstructured covariance matrix. Since linear mixed-models analysis can handle incomplete data, no procedure for imputation of missing data was utilized in the analysis [92]. For the analysis of BDI-II during the treatment phase, random intercepts across participants were estimated, and BDI-II's from every treatment session were included for the intervention group. Time was coded as 0 for baseline and as number of weeks from baseline for all subsequent measures. To control for differences in treatment duration, the time frame was made comparable

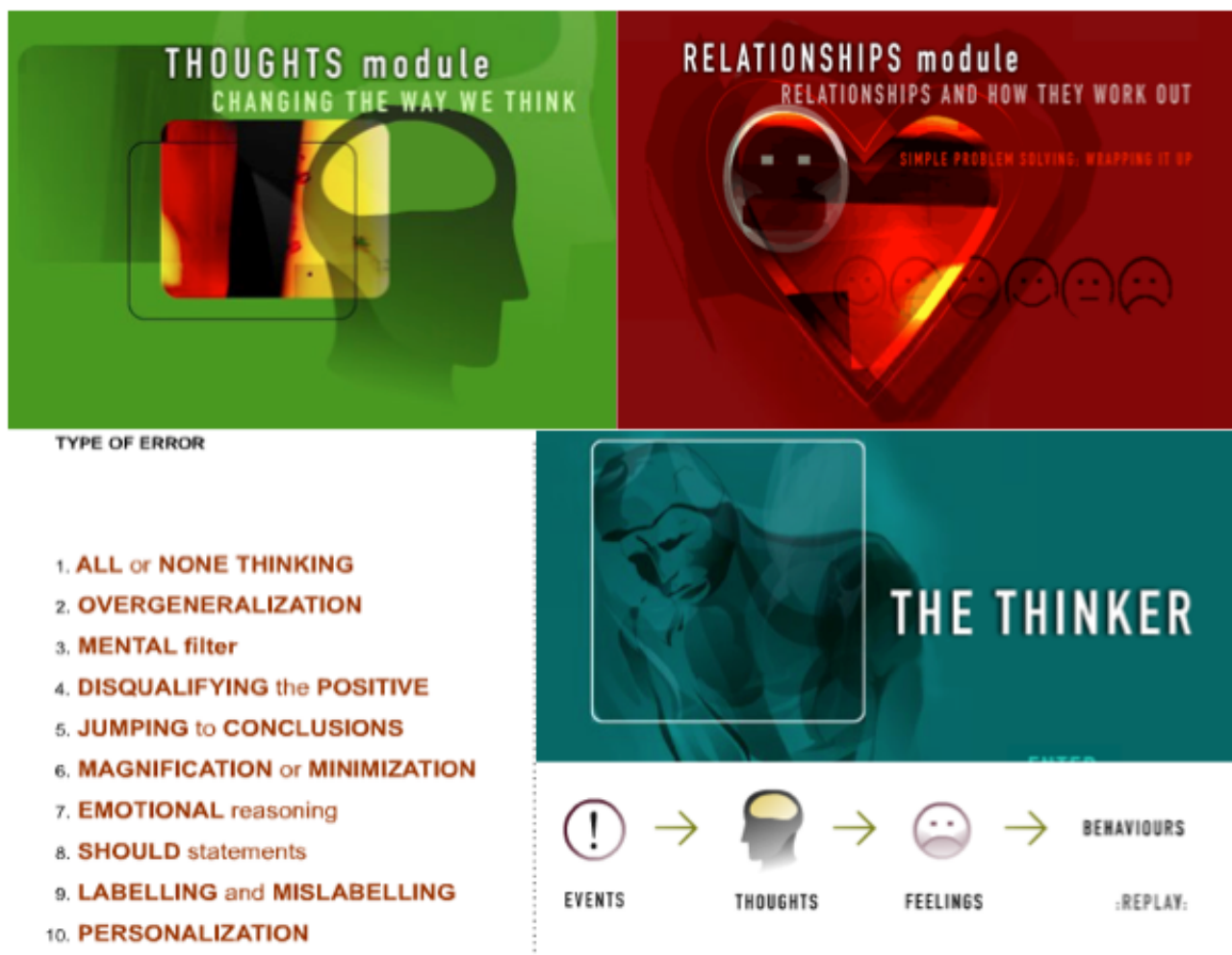
between groups by including only measures up to 7 weeks after baseline (the planned time frame for completing the intervention) for the intervention group in the main analysis. For the secondary measures and for the analysis including the 6-month follow-up data on the BDI-II, repeated measures linear mixed-models analysis was performed with occasion (baseline, posttest, 6-month follow-up) as the repeated factor. This procedure was deemed acceptable because linear regression analyses did not find treatment duration to be a significant predictor of symptom change during the treatment phase of the intervention group for any of the secondary measures (beta=-.16 to .28, $t_{29,35}=-0.91$ to 1.58, $P=.12-.92$). Scores on the last BDI-II from participants completing 5 or more weeks of treatment but missing formal posttest data (n=8), were included in the analysis, because this was considered to give a more accurate estimate of change over time.

For completers, analyses of covariance (ANCOVA) were performed with postsymptom scores as the dependent variable and preintervention symptom scores and treatment duration as covariates. Effect sizes (Cohen's *d*) were calculated for within- and between-group changes based on estimated means or beta

coefficients [93]. For the ITT analyses, calculations were based on pooled standard deviations calculated from the square root of each group's variance parameters from the mixed models analysis: the single variance estimate of the procedure with a random intercept, and the sum of the variance estimates at each time point of interest minus 2 times the covariance estimate between these time points for the repeated procedure [94,95]. For the completer analyses, the square root of the mean square error, equivalent to the pooled standard deviation, was used. A Cohen's *d* of 0.2 was considered a small effect, 0.5 a medium effect, and 0.8 or more a large effect [96].

Clinically significant changes on the BDI-II were assessed using the criteria for reliable change and cutoff points developed for the BDI by Seggar et al [97], based on the definition by Jacobson and Truax [98]. Recovery was defined as the combination of reliable improvement (a change of more than 8.5 points on the BDI-II) and endpoint symptom level below the clinical cutoff of 14.3. For those with subclinical symptoms at baseline, reliable improvement required a change of more than 4.6 points and recovery required reliable improvement plus an endpoint less than 4.1 [97]. Change scores between baseline and 7 weeks of treatment were used for the intervention group.

Figure 1. Screenshots from MoodGYM.



Results

Sample Characteristics

Figure 2 shows the flow of participants through the trial. Of the 128 individuals screened for participation, 106 (83%) were found eligible. Most participants, including 49 of 52 (94%) in the intervention group and 48 of 54 (89%) in the delayed-treatment group, were recruited from GPs. The remaining participants were recruited from waitlists at 2 outpatient clinics (n=3), from primary care nurses (n=4), and from a low-threshold clinic for youth (n=2). Postintervention measures were completed by 71% (37/52) and 87% (47/54) of the participants in the intervention and the delayed-treatment group, respectively. The 6-month follow-up assessment was completed by 42 participants (81%) in the intervention group and by 34 participants (63%) in the control group.

Group and educational level emerged as significant predictors of dropout at postintervention. The odds of dropping out before the posttest was significantly higher for participants in the intervention group relative to participants in the

delayed-treatment group (OR 3.03, 95% CI 1.08-8.47, $P=.04$), and significantly lower for participants with higher education relative to participants with a lower educational level (OR 0.36, 95% CI 0.13-0.99, $P=.048$). No other demographic or clinical variables predicted dropout at postintervention.

Table 1 shows the descriptive characteristics of the sample. For most demographic and clinical variables the 2 groups did not differ significantly at baseline ($P=.10-.90$). However, the groups differed significantly with regard to age ($P=.045$), with the intervention group being slightly older than the control group. The groups also differed on the variable comorbid anxiety ($P=.03$), with the number of participants with an anxiety disorder being significantly higher in the control group compared to the intervention group. Baseline scores on all symptom and outcome measures were comparable between the groups ($P=.13-.87$).

Further exploration of the 2 groups with regard to anxiety level as measured with BAI shows that although a higher proportion of the control group had symptoms corresponding to moderate or severe anxiety compared to the intervention group, the distribution across the categories minimal, mild, moderate, and severe anxiety did not differ significantly ($P=.29$).

Figure 2. Flow of participants through the trial.

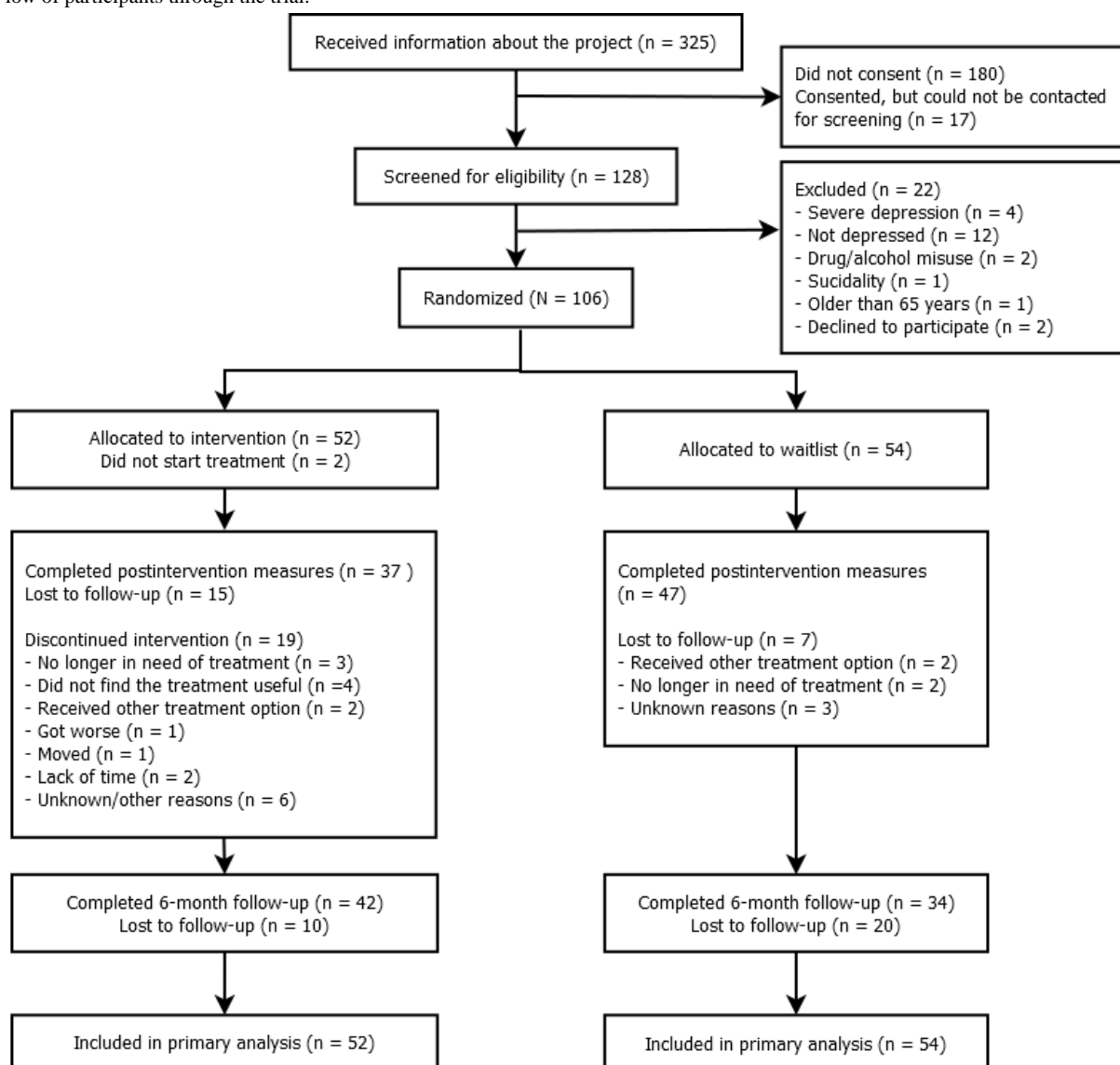


Table 1. Participant characteristics at baseline.

Participant characteristics	Intervention (n=52)	Waitlist (n=54)	Total (N=106)
Gender, n (%)			
Male	15 (28.8)	14 (25.9)	29 (27.4)
Female	37 (71.2)	40 (74.1)	77 (72.6)
Age (years)			
Mean (SD)	38.3 (12.2)	33.9 (9.9)	36.1 (11.3) ^a
Range	19 - 63	18 - 58	18 - 63
Marital status, n (%)			
Married/living together	28 (53.8)	31 (57.4)	59 (55.7)
Separate living	2 (3.8)	4 (7.4)	6 (5.7)
Divorced	3 (5.8)	5 (9.3)	8 (7.5)
Single	19 (36.5)	14 (25.9)	33 (31.1)
Highest educational level, ^b n (%)			
Compulsory school (9 or 10 years)	6 (11.5)	1 (1.9)	7 (6.6)
High school	16 (30.8)	26 (48.1)	42 (39.6)
University, 3 years	17 (32.7)	14 (25.9)	31 (29.2)
University, ≥5 years	13 (25.0)	12 (22.2)	25 (23.6)
Employment, n (%)			
Employed (full- or part-time)	37 (71.2)	37 (68.5)	74 (69.8)
Student	5 (9.6)	6 (11.1)	11 (10.4)
Long term sick	7 (13.5)	3 (5.6)	10 (9.4)
Homemaker	1 (1.9)	5 (9.3)	6 (5.7)
Unemployed	2 (3.8)	3 (5.6)	5 (4.7)
Sick leave (employed sample)	18 (48.6)	21 (56.8)	39 (52.7) ^c
Present treatment, ^d n (%)			
Medication	13 (25.0)	7 (13.0)	20 (18.9)
Other treatment	7 (13.5)	4 (7.4)	11 (10.4)
None	36 (69.2)	44 (81.5)	80 (75.5)
Treatment history, ^e n (%)			
Earlier	32 (61.5)	30 (55.6)	62 (58.5)
None	18 (34.6)	23 (42.6)	41 (38.7)
Depression current, ^f n (%)	26 (50.0)	25 (46.3)	51 (48.1)
Number of major depressive episodes, ^g n (%)			
0	5 (9.6)	6 (11.1)	11 (10.4)
1	16 (30.8)	18 (33.3)	34 (32.1)
2 - 4	14 (26.9)	15 (27.8)	29 (27.4)
≥5	14 (26.9)	11 (20.4)	25 (23.6)
Comorbidity, ^h n (%)			
Anxiety	12 (23.1)	23 (42.6)	35 (33.0) ^a
Other axis-I disorder	1 (1.9)	4 (7.4 %)	5 (4.7)

Participant characteristics	Intervention (n=52)	Waitlist (n=54)	Total (N=106)
Symptom measures, ⁱ mean (SD)			
AUDIT	4.8 (3.8)	5.4 (4.3)	5.1 (4.0)
DUDIT	0.4 (2.2)	0.7 (2.5)	0.6 (2.3)
Internet Use, ^j n (%)			
Daily	50 (96.2)	42 (77.8)	92 (86.8)
Weekly	0 (0.0)	3 (5.6)	3 (2.8)

^a $P < .05$.

^b0.9% missing.

^c36.8% of the total sample.

^dMedication is antidepressants; other treatment is psychological therapy other than CBT.

^e2.8% missing.

^fMajor depressive episode fulfilling *DSM-IV* criteria.

^g6.6% missing.

^hAnxiety includes panic disorder, agoraphobia, social phobia, and generalized anxiety; other axis-1 disorders include bipolar disorder, obsessive-compulsive disorder, bulimia, and posttraumatic stress disorder.

ⁱAUDIT: Alcohol Use Disorders Identification Test, 0.9% missing; DUDIT: Drug Use Disorders Identification Test, 0.9% missing.

^j10.4% missing.

Attrition and Adherence

Of the 52 participants in the intervention group, 31 (60%) adhered to the treatment program in that they completed MoodGYM and attended at least 7 sessions. Total nonadherence was 40% (21/52). Reasons for nonadherence are summarized in Figure 2. Overall, the sample starting treatment (n=50) completed a mean of 3.8 (SD 1.7) of the 5 modules, attended a mean of 7.2 (SD 2.3) sessions, with average session duration of 27.7 (SD 6.2) minutes. The average number of weeks in treatment was 11.3 (SD 7.2). Total time spent by therapists ranged from 70 to 506 minutes (mean 242.1, SD 96.6). A previous study found that versions of the MoodGYM program encompassing Module 2 (extended CBT) were associated with greater improvements than versions excluding this module [91]. This suggests that Module 2 may be a particularly important treatment component. Of the 50 participants starting treatment, 86% (n=43) completed 2 or more modules, indicating that they may have completed enough of the treatment program to generate beneficial outcomes.

Depression and Anxiety Symptoms

Table 2 depicts the preintervention, postintervention, and 6-month follow-up means and standard deviations for each group, as well as within-group and between-group effect sizes. The ITT analysis for the primary outcome measure, BDI-II, revealed a significant time by treatment group interaction ($F_{1,244.83}=9.55$, $P=.002$, $d=0.65$). There was also a significant effect of time ($F_{1,245.37}=11.87$, $P=.001$). Both groups experienced significant improvements of depressive symptoms during the intervention phase, but this improvement was significantly larger in the intervention group compared to the delayed-treatment group. Because most of the intervention group had not yet completed treatment at 7 weeks, the analysis was repeated using scores from the intervention group up to 8, 10, 12, and 14 weeks.

The interaction term remained significant ($F_{1,268.81-333.62}=6.34-10.88$, $P=.01-.001$, $d=0.54-0.67$). Repeated linear mixed-models analysis also found significant different time trends for the groups between posttest and 6-month follow-up ($t_{81.17}=2.88$, $P=.005$). During this time, the delayed-treatment group had received treatment, and pairwise comparisons indicated a significant decrease in symptoms in this group ($P=.001$), whereas level of depressive symptoms in the intervention group remained stable ($P=.56$).

The ITT analysis for the BAI revealed no significant interaction between treatment group and occasion ($F_{1,84.31}=0.37$, $P=.69$). Pairwise comparisons showed that both groups improved significantly between baseline and posttest ($P=.007$ and $P=.02$ for the intervention and control group, respectively; see Table 2). Between posttest and 6-month follow-up, the control group further improved ($P=.04$), whereas the intervention group remained unchanged ($P=.36$). Analyses for the HADS subscales yielded significant group by occasion interactions for both the depression subscale ($F_{1,78.05}=14.68$, $P<.001$) and the anxiety subscale ($F_{1,78.07}=8.10$, $P=.001$). Pairwise comparisons for both subscales found that the intervention group, but not the control group, reduced their scores significantly between pretest and posttest ($P<.001$ and $P=.46$, respectively, for the depression subscale and $P<.001$ and $P=.44$, respectively, for the anxiety subscale; Table 2). Between posttest and 6-month follow-up, the control group experienced a significant reduction in both depressive and anxiety symptoms ($P=.001$ and $P<.001$, respectively), whereas scores did not change significantly for the intervention group for the depressive subscale ($P=.07$) or for the anxiety subscale ($P=.80$).

For all reported ITT and modified ITT analyses, baseline anxiety symptoms on the BAI and age were included as covariates, with the exception that baseline BAI scores were not included as a covariate in the analysis of BAI. The effect of BAI scores was

consistently significant ($P=.02$ to $P<.001$). The effect of age was significant in 2 of 8 models ($P=.008-.03$).

Completer Analyses

The results of the ANCOVA revealed the same pattern of results. There was a significant effect of group on posttest level of depressive symptoms measured with BDI-II ($n=84$; $P<.001$, $d=1.09$) and with HADS ($n=76$; $P=.002$, $d=0.95$), and level of anxious worry as measured with HADS ($n=76$; $P=.002$, $d=0.97$) after controlling for the effect of baseline symptoms and treatment duration. The results were not significant for anxiety as measured with BAI ($n=83$; $P=.47$, $d=0.20$). Thus, for depression and anxiety measured with HADS, the groups differed significantly in posttest scores after controlling for differences in preintervention scores and treatment duration. Baseline symptoms were significantly related to posttest symptoms for all measures ($P<.001$), whereas treatment duration did not significantly affect outcome ($P=.19-.92$).

Quality of Life

For the SWLS, the modified ITT analysis showed a significant interaction between treatment group and occasion ($F_{1,75.75}=8.49$, $P<.001$). Pairwise comparisons found a significant increase in satisfaction for the intervention group from pretest to posttest ($P<.001$), but no such change for the delayed-treatment group ($P=.52$; see Table 2). Between posttest and 6-month follow-up, there was a significant increase in life satisfaction in the control group ($P=.01$), whereas the intervention group did not experience significant changes ($P=.08$). The analyses for EQ Index yielded an overall significant difference in time trends between the groups ($F_{1,67.42}=3.55$, $P=.03$). Between pretest and posttest, there was no significant interaction between treatment group and occasion ($t_{68.82}=-1.00$, $P=.32$), with both groups improving at comparable rates over time. However, there was a significant group by occasion interaction between posttest and 6-month follow-up ($t_{63.96}=-2.66$, $P=.01$, with the control group showing significant improvement ($P=.02$), whereas there were no significant changes in the intervention group ($P=.18$). For the EQ VAS, the overall interaction between group and occasion was not significant ($F_{1,71.32}=2.25$, $P=.11$). Between pretest and posttest both the intervention group ($P<.001$) and the control group ($P=.03$) experienced significantly improved self-reported

health state (Table 2), but the interaction between group and occasion was not significant ($t_{66.07}=-1.94$, $P=.06$). Between posttest and 6-month follow-up, pairwise comparisons showed a slight, but nonsignificant, improvement in the control group ($P=.06$), whereas scores in the intervention group remained stable ($P=.74$).

Completer Analyses

Similar results were found for the completer analyses, in which ANCOVA showed a significant effect of group for the SWLS ($n=66$) after controlling for baseline levels of life satisfaction and treatment duration ($P=.006$, $d=0.86$). There was no significant effect of group on health-related quality of life at posttest, EQ Index ($n=63$; $P=.56$, $d=0.18$), or health state at posttest, EQ VAS ($n=61$; $P=.11$, $d=0.52$), despite a moderate effect size for the latter. The effect of baseline scores was significant for all measures ($P\le.002$). There were no significant effects of treatment duration ($P=.14-.99$).

Clinical Significance of Changes in Depressive Symptoms

Table 3 presents data on clinically significant change on the BDI-II, based on scores at 7 weeks for the intervention group. The ITT procedure was used (classifying all who did not start treatment or did not complete their waitlist period as nonresponders). The results of the chi-square tests for the full sample and the sample with BDI scores above clinical cutoff show that significantly more participants recovered in the intervention group compared to the delayed-treatment control group. Conversely, a significantly smaller proportion of the intervention group experienced no change within the intervention period. For the sample fulfilling the criteria for a major depressive episode at baseline, the same trend was evident, but the difference in rate of recovery did not reach significance. The rates of improvement and deterioration were similar in the 2 groups for all analyses. The same analysis carried out after excluding participants ($n=8$) who during the waitlist or intervention period started or increased their dosage of antidepressant medication, or commenced other psychological treatment, produced similar patterns of results. At 7 weeks of treatment 37 of 52 participants (71%) in the intervention group had completed 2 or more modules, whereas 15% ($n=8$) had completed treatment.

Table 2. Estimated means (EM),^a observed means (OM), observed standard deviations (SD), standard deviations based on linear mixed-models variance estimates (SDm), and effect sizes from pretreatment (pre) to posttreatment (post) and pretreatment to 6-month follow-up (6 m) for the intervention and the delayed-treatment control group.

Measures	Intervention (n=52)				Delayed treatment (n=54)				Effect size (Cohen's <i>d</i>) ^b					
	EM	OM	SD	SDm	EM	OM	SD	SDm	Pre-post			Pre-6 m		
									B	Wi	Wdt	B	Wi	Wdt
BDI-II									0.65	-0.98	-0.65	-0.12	-0.81	-1.02
Pretreatment	21.37	21.13	6.85		21.85	22.27	6.74							
7 weeks	15.15 ^c	14.20 ^c	8.15	5.86	19.07	18.63	8.64	4.63						
6 months	13.39	12.45	9.32	10.35	11.86	12.82	10.98	9.56						
BAI									0.08	-0.41	-0.35	-0.12	-0.48	-0.65
Pretreatment	12.23	12.05	11.10		15.10	15.33	10.90							
Posttreatment	8.80	8.36	9.26	8.21	12.30	12.83	8.10	8.03						
6 months	7.46	7.07	6.61	10.37	9.19	9.41	9.79	9.05						
HADS Depression									1.10	-1.17	-0.11	-0.10	-0.53	-0.65
Pretreatment	8.21	8.08	2.92		7.40	7.61	3.13							
Posttreatment	4.67	4.24	2.61	3.13	7.07	7.19	3.63	2.70						
6 months	5.91	5.76	4.05	4.29	4.66	4.85	3.77	4.16						
HADS Anxiety									0.74	-0.60	0.13	-0.13	-0.52	-0.55
Pretreatment	9.14	8.81	3.95		9.15	9.59	4.59							
Posttreatment	7.15	6.74	3.69	3.29	9.52	10.07	4.18	3.04						
6 months	6.99	6.57	4.16	3.99	6.40	6.94	4.79	4.88						
SWLS									0.85	0.85	0.12	-0.12	0.35	0.52
Pretreatment	16.41	16.54	5.25		16.83	16.36	5.75							
Posttreatment	20.38	21.46	6.04	4.60	17.28	17.21	5.02	3.69						
6 months	18.66	19.00	6.63	6.42	19.79	20.00	6.91	5.52						
EQ-5D VAS									0.47	0.79	0.42	-0.01	0.71	0.55
Pretreatment	58.59	59.13	18.68		56.01	54.93	17.87							
Posttreatment	71.13	73.88	15.30	16.35	61.67	60.77	17.09	13.10						
6 months	70.12	71.79	14.65	16.04	67.78	66.41	20.86	21.56						
EQ-5D Index									0.22	0.59	0.26	-0.33	0.26	0.74
Pretreatment	0.63	0.63	0.23		0.61	0.60	0.25							
Posttreatment	0.75	0.80	0.17	0.24	0.68	0.67	0.24	0.27						
6 months	0.70	0.72	0.26	0.30	0.78	0.77	0.20	0.21						

^aEstimated means (except for BAI) are adjusted for the covariates baseline BAI score and age.

^bB: between-group effect size; Wi: within-group effect size for the intervention group; Wdt=within-group effect size for the delayed-treatment group.

^cEstimated mean after completing treatment=12.43, observed mean after completing treatment=11.34.

Table 3. Proportion of participants reaching the criteria for clinically significant improvement on the Beck Depression Inventory-II (BDI-II) at 7 weeks of treatment and results of chi-square tests (χ^2).

Treatment response	Full sample, n (%) (N=106)		Baseline BDI-II above clinical cutoff, ^a n (%) (n=90)		Current major depressive episode diagnosis, n (%) (n=51)				
	Intervention (n=52)	Control (n=54)	χ^2_1	Intervention (n=42)	Control (n=48)	χ^2_1	Intervention (n=26)	Control (n=25)	χ^2_1
Recovered	17 (32.7)	5 (9.3)	8.8 ^b	15 (35.7)	5 (10.4)	8.3 ^b	8 (30.8)	3 (12.0)	2.7
Improved	8 (15.4)	5 (9.3)	0.9	5 (11.9)	4 (8.3)	0.3	6 (23.1)	3 (12.0)	1.1
No change	26 (50.0)	41 (75.9)	7.7 ^b	21 (50.0)	38 (79.2)	8.4 ^b	11 (42.3)	18 (72.0)	4.6 ^c
Deteriorated	1 (1.9)	3 (5.6)	0.3	1 (2.4)	1 (2.1)	0.01	1 (3.8)	1 (4.0)	0.001

^aBDI-II>14.^b $P<.01$.^c $P<.05$.

Therapist Effects

Therapist effects were investigated by looking at the interaction between therapist and time for the intervention group. The analyses indicated a significant difference between the 2 therapists when analyzing BDI-II scores up to 7 and 8 weeks of treatment ($P=.03-.04$). This effect no longer reached significance when including scores up to 10, 12, and 14 weeks of treatment ($P=.05-.32$). The analyses did not yield differential treatment effects across the 2 therapists for the HADS depression subscale ($P=.87$), nor for any other outcome measure ($P=.50-.94$). An exploratory linear regression analysis showed that symptom change in the intervention group was not significantly predicted by the total time spent by the therapists for any outcome measure (beta=-.12 to .26, $t_{29,48}=-0.81$ to 1.52, $P=.14-.96$).

Treatment Satisfaction

Table 4 shows the response frequencies for questions regarding satisfaction with the treatment. The results are reported for participants in both groups (intervention group: n=39; delayed-treatment control group: n=26) who completed the full treatment or parts of it. Overall satisfaction with the treatment was high, with 89% (58/65) giving the intervention as a whole a rating of 4 or 5 on a scale with 5 being very satisfied (see Table 4). Most participants also indicated that they would recommend the combined intervention to a friend with a similar problem. The ratings of the MoodGYM program were positive,

but somewhat more moderate with between 50% and 60% giving clearly positive ratings (4 or 5 on the 5-point scales, see Table 4) to the benefit of the program, the usefulness of the exercises, and the relevance of the thematic content, and none rating the program as not useful or relevant. The benefit of the treatment sessions and the relationship with the therapist were rated positively by more than 90% (60/65 and 64/65, respectively) of the participants.

Service Use and Work Status After Treatment

Of the 76 participants who completed the follow-up assessment, 45% (19/42) of participants in the intervention group and 38% (13/34) of participants in the control group had received treatment for mental health problems during the 6-month follow-up period. Two participants (3%) had been hospitalized, 19 (25%) had used antidepressant medication (15 currently using), 26 (34%) had received psychological treatment individually or group therapy, and 16 (21%) had received treatment from their GP. Of the 19 participants reporting use of antidepressants, only 6 had commenced this treatment during the follow-up period.

With regard to work status, 6 of 42 respondents (14%) in the intervention group reported that they had been on sick leave at some point in the follow-up period due to feeling tired, stressed, or experiencing mental health problems, whereas 9 of 34 respondents (26%) in the control group reported sick leave during this period.

Table 4. Response frequencies regarding satisfaction and experiences with the treatment (n=65).

Item	Satisfaction/experience scale				
	1	2	3	4	5
Overall satisfaction with the treatment					
Scale	Very dissatisfied				Very satisfied
n (%)	0 (0.0)	0 (0.0)	7 (10.8)	40 (61.5)	18 (27.7)
Change in symptoms					
Scale	Much worse		Neither nor	Much improved	
n (%)	0 (0.0)	2 (3.1)	16 (24.6)	26 (40.0)	21 (32.3)
Would recommend the treatment to a friend with a similar problem					
Scale	Definitely not				Yes, definitely
n (%)	0 (0.0)	0 (0.0)	4 (6.2)	21 (32.3)	40 (61.5)
Benefit of using MoodGYM					
Scale	No benefit				Highly beneficial
n (%)	0 (0.0)	4 (6.2)	24 (36.9)	32 (49.2)	5 (7.7)
The usefulness of the exercises in MoodGYM					
Scale	Not useful				Very useful
n (%)	0 (0.0)	5 (7.7)	23 (35.4)	28 (43.1)	9 (13.8)
Relevance of the thematic content of MoodGYM ^a					
Scale	No relevance				Highly relevant
n (%)	0 (0.0)	I: 4 (10.3) C: 1 (3.8)	I: 20 (51.3) C: 6 (23.1)	I: 10 (25.6) C: 13 (50.0)	I: 5 (12.8) C: 6 (23.1)
Benefit of the follow-up sessions					
Scale	No benefit				Highly beneficial
n (%)	0 (0.0)	0 (0.0)	5 (7.7)	34 (52.3)	26 (40.0)
Satisfaction with the number of sessions					
Scale	Too few		Just enough	Too many	
n (%)	1 (1.5)	6 (9.2)	55 (84.6)	3 (4.6)	0 (0.0)
Relationship to the therapist					
Scale	Very negative				Very positive
n (%)	0 (0.0)	0 (0.0)	1 (1.5)	14 (21.5)	50 (76.9)

^a $P < .05$, frequencies reported separately for the intervention (I) and delayed-treatment control (C) groups.

Discussion

Principal Findings

The results of the present study indicate that a guided self-help intervention combining the MoodGYM program with face-to-face therapist support can be effective in reducing depressive symptoms for a sample of mildly to moderately depressed individuals recruited from primary care. The intervention also had significant positive effects on symptoms of anxious worry, and participants experienced significant improvements in global satisfaction with life. At 6-month follow-up, positive gains in terms of reduction of depressive and anxious symptoms were largely maintained, whereas improvements in life satisfaction were partly maintained. The

rate of nonadherence (40%) was moderate and the evaluations of the treatment as a whole were predominately positive.

These findings are consistent with previous research in which favorable outcomes have been shown for treatments combining MoodGYM and face-to-face support from a professional [29,37]. The trials are not fully comparable, though, because the present study used a delayed-treatment control condition, whereas both previous trials used comparison groups that received more active treatments. This makes direct comparisons of between-group effect sizes difficult. However, the effect of guided self-help for mild to moderate depression using other Internet-based programs has generally been in the moderate to large range [19,27]. The magnitude of the between-group effect size on depression measured with BDI-II in this study was within this range, but somewhat below average, whereas the effects on

SWLS and on depression and anxiety measured with HADS were commensurable to effects found previously for similar measures [12-14,27,99,100]. Also, the rates of recovery and of improvement and recovery combined (33% and 48%, respectively) are in good accordance with results of prior investigations, in which clinically significant improvement and recovery have varied between 25% and 50% [12,14,15,30,100]. The nature of the guidance provided in these studies is somewhat heterogeneous, with some studies, such as the present study, defining guidance as active engagement in the therapeutic process, whereas several other studies have focused primarily on providing feedback and encouragement. However, there is no indication of significantly differential treatment effects depending on the nature of the guidance to date [19,100]. The somewhat smaller between-group differences of the present study must be seen in relation to the relatively high degree of positive change in the control group, with effect sizes being in the small to moderate range for several measures (see Table 2). Almost half of the sample did not fulfill the criteria for a major depressive episode on enrollment in the trial, and previous studies of minor depression have shown high rates of placebo response in primary care patients [101] and substantial likelihood of spontaneous remission in the general population [102]. In the present study, the control group was also free to access usual treatment in general practice during the waiting period. In addition, prior to entering the study control participants participated in a screening session in which they had the opportunity to describe their problems, something that could have a therapeutic effect per se. These factors may partly explain the positive gains in the delayed-treatment group and, hence, the modest differences in outcome between the groups.

The results of the 6-month follow-up are also encouraging in that improvements of depression and anxiety symptoms were largely maintained, and the number of participants reporting sick leave due to mental health problems was substantially smaller during follow-up compared to baseline (20% and 53%, respectively). The proportion of participants using antidepressant medication at baseline and during the follow-up period was comparable, but there was a considerable increase in the number of participants who had accessed psychological treatment during follow-up (34%) compared to baseline (10%). This is, however, not surprising because many participants had already been referred to such treatment when entering the project, but may still have contributed to further improvements and maintenance of symptom reduction. For some measures, particularly the depression subscale of the HADS and the SWLS, there was a tendency toward an increase in symptoms and lowering of life satisfaction during the follow-up period. The inclusion of booster sessions after the completion of the active treatment phase could be a measure to accomplish continued use of helpful techniques and skills and the prevention of symptom relapse. Further research is needed to clarify this issue.

The results of the present trial are also consistent with research suggesting that MoodGYM, despite its main focus on depressive thought content, can have significant positive effects on anxiety symptoms [29,35,103]. In the present study, significant treatment effects were found for anxiety symptoms measured with HADS, but not with BAI. The questions on the HADS focus primarily

on symptoms such as worry, nervousness, and not being able to relax. This is in good accordance with the core symptoms of the Goldberg Anxiety Scale [104], which has been used in several studies with MoodGYM. In comparison, the BAI incorporates both these subjective anxiety symptoms, as well as 3 more somatic symptom clusters: neurophysiological, autonomic, and panic [63]. The MoodGYM program focuses primarily on restructuring dysfunctional thinking and does not include an introduction to the physiology of anxiety or other treatment techniques for anxiety. Thus, the present results, with the program showing effects on anxious worry but not on more physiological symptoms, seem to be in-line with the thematic focus of the program.

The adherence rate in the present study is comparable to that observed in other online guided interventions, in which adherence varied between 55% and 75% [13,14,23,30,99]. The rates of adherence are also comparable to those seen in other psychotherapy research and in regular clinical practice [105-107]. This level of adherence in the current study, and the high proportion (89%) of completers reporting being satisfied with the treatment, points to the acceptability of the intervention. The evaluation of the MoodGYM program was somewhat more moderate with between 50% and 60% allocating an unambiguous positive rating to the benefit and relevance of the program. This moderate level of satisfaction in the present adult sample may arise from the fact that the MoodGYM program was originally targeted at youth and young adults. Although the therapists emphasized the applicability of the principles for all age groups when introducing the intervention, and many participants managed well to make use of the content, participants frequently characterized the program as “too young.” As these aspects of treatment were not formally measured, these factors require further investigation to be properly elucidated.

Strengths and Limitations

The current study was designed to trial a treatment procedure prior to its evaluation in general practice. Therefore, we sought to ensure a high level of internal validity while at the same time aiming to increase external validity by reflecting the heterogeneity of patients in real clinical practice. One strength of the study is the relatively heterogeneous sample of participants with regard to the range of depression and anxiety symptoms. There was also substantial comorbidity with anxiety disorders, although lower than rates found in population-based studies [108,109]. The fact that 83% of screened participants were found eligible also indicates that the sample is representative of those who opted for this choice of treatment.

An overarching focus in designing the intervention was feasibility for implementation in general practice. Studies have suggested that GPs may find the implementation of CBT techniques too time-consuming [49,110]. Therefore, sessions were primarily supportive and structured by the Web-based program. To allow for flexibility and increase feasibility, a guideline script rather than a more comprehensive manual guided each consultation. This lack of rigid standardization may have introduced some variability in treatment fidelity.

Blinding of participants was not possible for obvious reasons in this trial. However, the control group was blinded to the status of the waitlist as a control condition. Waitlists for treatment is the norm in Norwegian mental health care, and the short wait for the present treatment compared to other treatment options, may have minimized negative effects (“nocebo effects”) in the control group.

The present study also has several limitations that need to be addressed. First, the design of the study with only 1 intervention group receiving a compound of intervention elements does not allow for tests of the specific contribution of MoodGYM and the face-to-face consultations. Second, the lack of allocation concealment and blinding, and the role of the first and second authors as therapists in the trial introduced a risk of bias that may have inflated the treatment effects. Unfortunately, resource constraints prevented the use of independent therapists. The use of self-report measures rather than therapist assessments does alleviate this problem to some degree. Biased outcome assessments were further prevented by ensuring that a research assistant without knowledge of the participants’ condition assignment collected posttest data. Third, the sole reliance on self-report is a limitation in itself. Independent preassessments and postassessments by a clinician blinded to condition allocation would have been preferable and would have strengthened the results. Fourth, at preintervention there was a lack of comparability in diagnosed anxiety between the groups, with a significantly larger proportion of the control group fulfilling the criteria for an anxiety disorder. Despite this difference in diagnosed anxiety, scores on 2 different anxiety scales (HADS and BAI) were not significantly different, which suggests comparable anxiety levels in the groups. To minimize the effect of differences in anxiety, all primary analyses were controlled for anxiety level as measured with BAI, for which the observed difference was most significant. Fifth, the use of an unequal number of assessments of depressive symptoms in the 2 groups, with the intervention group having weekly assessments and the control group only completing a pretest and posttest, may have resulted in more favorable effects in the intervention group because of measurement effects. Previous studies of nonclinical samples have indicated that scores on the BDI tend to decrease with repeated administration [111,112]. Whether this holds for clinical samples is less certain. In this study, the effect of repeated measurement cannot be clearly distinguished from the treatment effect. However, comparable effects were also found for symptoms assessed only preintervention and postintervention in both groups. This indicates that the treatment had beneficial effects over and above possible measurement effects. Still, in light of this limitation, the results must be interpreted with caution. Sixth, the use of different administration formats for the assessments of the treatment phase and follow-up, (paper-and-pencil vs online questionnaires, respectively) can potentially introduce measurement bias. Although the 2 formats correlate highly, a previous study reported a significant difference in mean scores on the BDI-II and BAI, which makes switching of formats problematic [59]. Despite this limitation, the results should not be considered weakened for most measures because the direction of differences has generally suggested that online versions tend to inflate estimations of symptom severity and lower ratings of

quality of life [59,71,76], with the exception of BAI, for which Carlbring et al [59] found that means on the online version were lower compared to the paper-and-pencil version. The reliability of the 6-month follow-up results for the BAI may, therefore, be limited. Seventh, the multiplicity of outcomes increases the risk of type I errors. However, the main findings of the present trial would still be significant when employing the Bonferroni correction. This indicates the robustness of the findings. Finally, although the heterogeneity of the sample and the recruitment from primary care is a strength, the generalizability of the results is uncertain because the sample was a self-selected group. Based on the notifications by the GPs when informing a patient of the study, the estimated uptake (meeting up for screening) was 39% (128 of 325 who received information), which is slightly greater than the median uptake for computerized CBT [55]. Considering the extra barriers imposed by the research activities, this uptake rate is relatively high and indicates the possible acceptability of this treatment among depressed primary care patients. It also strengthens the generalizability of the results, by indicating that the self-selected group may be representative of a considerable proportion of the targeted group of primary care patients.

Potential Clinical Implications and Further Research

The positive treatment effects found for the intervention in the present study are encouraging and suggest that this intervention may have a potential for use in a stepped-care approach. The demand for mental health treatment is higher than what can be met by the current number of trained clinicians [113]. To increase availability of treatment, beneficial interventions must be delivered as efficiently as possible to as many people as possible. The present intervention is time-limited, and because the CBT elements are largely delivered by the program, primary care therapists with some training in CBT and MoodGYM, should be able to provide adequate guidance. In fact, studies show that guidance may be delivered effectively not only by trained clinicians, but also by mental health workers with limited experience and by computer technicians [114,115]. Thus, dissemination of the current intervention to regular primary care could be a step toward increasing access to psychological therapies. However, the moderate ratings of the benefit and relevance of the content of the Web-based program by an adult population points to the need for a variety of Web-based tools to make such treatments acceptable for a wider audience.

For practical reasons, we chose to use psychologists for this first evaluation. Therefore, further research is needed to determine if the present intervention would be as effective and acceptable in regular clinical practice when delivered by GPs or other primary care therapists. It may also be noted that the present intervention was more time-intensive than most other guided self-help interventions. However, since the role of the clinician was mainly supportive and facilitative and the main therapeutic input was delivered through a standardized treatment package, the intervention was regarded conceptually as a guided self-help intervention [31]. Similar effects have been found for low- and high-intensity guided Internet-based psychotherapy [19]. Further research should investigate if the present intervention with more limited therapist support could yield similar effects.

Conclusion

Despite its limitations, the present study indicates that an intervention combining the MoodGYM program with therapist support can be an effective treatment of depression in a sample of primary care patients. The intervention not only alleviates depressive symptoms, but also has positive and significant effects on symptoms of anxious worry and global satisfaction with life. Positive gains in terms of reduction of depressive and anxious symptoms were largely maintained at 6-month follow-up, and improvements in life satisfaction were partly maintained. Moderate rates of nonadherence and predominately

positive evaluations of the treatment as a whole also indicates the acceptability of the intervention. The intervention was designed to be suitable for implementation in primary health care, and could have a potential for use in a stepped-care approach. However, further research is necessary to determine whether it is equally effective when delivered in regular primary health care and whether the inclusion of booster sessions could further improve symptom maintenance. Further research is also needed to investigate whether the intervention is truly acceptable for the wider group of primary care patients and whether it is considered feasible and acceptable by GPs or other primary care therapists.

Acknowledgments

This study was funded by the Research Council of Norway (196423/V50). MoodGYM was provided by the Australian National University. The authors would like to acknowledge the contribution of Anthony Bennett, IT manager of the e-hub, the e-mental health research and development group at the Centre for Mental Health Research at ANU; Kylie Bennett, Development Manager at e-hub, who provided technical expertise in administration of the trial; and Ada Tam, e-hub Project Officer at ANU, who provided technical support for users in the trial. We would also like to thank Jan Abel Olsen at the Institute of Community Medicine at the University of Tromsø for contributing with his expertise on the EQ-5D scale. Last, but not least, we would like to give a special thanks to those participating in the trial.

Conflicts of Interest

KG is one of the authors of the MoodGYM intervention evaluated in the trial, and works for The Australian National University who provides free access to the MoodGYM program. KW and ME contributed in the process of translating MoodGYM into Norwegian. RSH, KL, TW, and NK have no financial or nonfinancial interests to declare in relation to this study.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [53].

[PDF File (Adobe PDF File), 999KB - [jmir_v15i8e153_app1.pdf](#)]

Multimedia Appendix 2

Trial Protocol as requested in CONSORT-EHEALTH [53].

[PDF File (Adobe PDF File), 232KB - [jmir_v15i8e153_app2.pdf](#)]

Multimedia Appendix 3

Trial information to participants as requested in CONSORT-EHEALTH [53].

[PDF File (Adobe PDF File), 115KB - [jmir_v15i8e153_app3.pdf](#)]

References

1. Ayuso-Mateos JL, Vázquez-Barquero JL, Dowrick C, Lehtinen V, Dalgard OS, Casey P, ODIN Group. Depressive disorders in Europe: prevalence figures from the ODIN study. *Br J Psychiatry* 2001 Oct;179:308-316 [FREE Full text] [Medline: [11581110](#)]
2. Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. *Eur Neuropsychopharmacol* 2005 Aug;15(4):357-376. [doi: [10.1016/j.euroneuro.2005.04.012](#)] [Medline: [15961293](#)]
3. Kringlen E, Torgersen S, Cramer V. A Norwegian psychiatric epidemiological study. *Am J Psychiatry* 2001 Jul;158(7):1091-1098. [Medline: [11431231](#)]
4. Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* 2007 Oct;6(3):168-176 [FREE Full text] [Medline: [18188442](#)]
5. Kilbourne AM, Daugherty B, Pincus HA. What do general medical guidelines say about depression care? Depression treatment recommendations in general medical practice guidelines. *Curr Opin Psychiatry* 2007 Nov;20(6):626-631. [doi: [10.1097/YCO.0b013e3282f0c4d3](#)] [Medline: [17921767](#)]

6. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011 Sep;21(9):655-679. [doi: [10.1016/j.euroneuro.2011.07.018](https://doi.org/10.1016/j.euroneuro.2011.07.018)] [Medline: [21896369](https://pubmed.ncbi.nlm.nih.gov/21896369/)]
7. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *J Consult Clin Psychol* 2008 Dec;76(6):909-922. [doi: [10.1037/a0013075](https://doi.org/10.1037/a0013075)] [Medline: [19045960](https://pubmed.ncbi.nlm.nih.gov/19045960/)]
8. Bebbington P, Brugha T, Meltzer H, Jenkins R, Ceresa C, Farrell M, et al. Neurotic disorders and the receipt of psychiatric treatment. *Int Rev Psychiatry* 2003;15(1-2):108-114. [doi: [10.1080/0954026021000046010](https://doi.org/10.1080/0954026021000046010)] [Medline: [12745317](https://pubmed.ncbi.nlm.nih.gov/12745317/)]
9. Kessler RC, Berglund PA, Bruce ML, Koch JR, Laska EM, Leaf PJ, et al. The prevalence and correlates of untreated serious mental illness. *Health Serv Res* 2001 Dec;36(6 Pt 1):987-1007 [FREE Full text] [Medline: [11775672](https://pubmed.ncbi.nlm.nih.gov/11775672/)]
10. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, et al. A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technol Assess* 2001;5(35):1-173 [FREE Full text] [Medline: [12387733](https://pubmed.ncbi.nlm.nih.gov/12387733/)]
11. 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/)]
12. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. *Cogn Behav Ther* 2009;38(4):206-221. [doi: [10.1080/16506070802408086](https://doi.org/10.1080/16506070802408086)] [Medline: [19221919](https://pubmed.ncbi.nlm.nih.gov/19221919/)]
13. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldö V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. *Br J Psychiatry* 2005 Nov;187:456-461 [FREE Full text] [doi: [10.1192/bjp.187.5.456](https://doi.org/10.1192/bjp.187.5.456)] [Medline: [16260822](https://pubmed.ncbi.nlm.nih.gov/16260822/)]
14. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. *Cogn Behav Ther* 2011;40(4):251-266. [doi: [10.1080/16506073.2011.616531](https://doi.org/10.1080/16506073.2011.616531)] [Medline: [22060248](https://pubmed.ncbi.nlm.nih.gov/22060248/)]
15. 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/)]
16. McCrone P, Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro DA, et al. Cost-effectiveness of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry* 2004 Jul;185:55-62 [FREE Full text] [Medline: [15231556](https://pubmed.ncbi.nlm.nih.gov/15231556/)]
17. Marks IM, Mataix-Cols D, Kenwright M, Cameron R, Hirsch S, Gega L. Pragmatic evaluation of computer-aided self-help for anxiety and depression. *Br J Psychiatry* 2003 Jul;183:57-65 [FREE Full text] [Medline: [12835245](https://pubmed.ncbi.nlm.nih.gov/12835245/)]
18. Marks IM, Kenwright M, McDonough M, Whittaker M, Mataix-Cols D. Saving clinicians' time by delegating routine aspects of therapy to a computer: a randomized controlled trial in phobia/panic disorder. *Psychol Med* 2004 Jan;34(1):9-17. [Medline: [14971623](https://pubmed.ncbi.nlm.nih.gov/14971623/)]
19. Titov N. Internet-delivered psychotherapy for depression in adults. *Curr Opin Psychiatry* 2011 Jan;24(1):18-23. [doi: [10.1097/YCO.0b013e32833ed18f](https://doi.org/10.1097/YCO.0b013e32833ed18f)] [Medline: [20827199](https://pubmed.ncbi.nlm.nih.gov/20827199/)]
20. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J Med Internet Res* 2004 Dec 22;6(4):e46 [FREE Full text] [doi: [10.2196/jmir.6.4.e46](https://doi.org/10.2196/jmir.6.4.e46)] [Medline: [15631970](https://pubmed.ncbi.nlm.nih.gov/15631970/)]
21. Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G. Self-guided psychological treatment for depressive symptoms: a meta-analysis. *PLoS One* 2011;6(6):e21274 [FREE Full text] [doi: [10.1371/journal.pone.0021274](https://doi.org/10.1371/journal.pone.0021274)] [Medline: [21712998](https://pubmed.ncbi.nlm.nih.gov/21712998/)]
22. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. *J Med Internet Res* 2009;11(2):e15 [FREE Full text] [doi: [10.2196/jmir.1151](https://doi.org/10.2196/jmir.1151)] [Medline: [19632969](https://pubmed.ncbi.nlm.nih.gov/19632969/)]
23. Clarke G, Kelleher C, Hornbrook M, Debar L, Dickerson J, Gullion C. Randomized effectiveness trial of an Internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. *Cogn Behav Ther* 2009;38(4):222-234 [FREE Full text] [doi: [10.1080/16506070802675353](https://doi.org/10.1080/16506070802675353)] [Medline: [19440896](https://pubmed.ncbi.nlm.nih.gov/19440896/)]
24. Farrer L, Christensen H, Griffiths KM, Mackinnon A. Internet-based CBT for depression with and without telephone tracking in a national helpline: randomised controlled trial. *PLoS One* 2011;6(11):e28099 [FREE Full text] [doi: [10.1371/journal.pone.0028099](https://doi.org/10.1371/journal.pone.0028099)] [Medline: [22140514](https://pubmed.ncbi.nlm.nih.gov/22140514/)]
25. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CE, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided Internet-based self-help intervention for the prevention of depression: a randomized controlled trial. *Clin Psychol Psychother* 2013;20(1):10-27. [doi: [10.1002/cpp.770](https://doi.org/10.1002/cpp.770)] [Medline: [21887811](https://pubmed.ncbi.nlm.nih.gov/21887811/)]
26. Spek V, Cuijpers P, Nyklíček I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007 Mar;37(3):319-328. [doi: [10.1017/S0033291706008944](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
27. Johansson R, Andersson G. Internet-based psychological treatments for depression. *Expert Rev Neurother* 2012 Jul;12(7):861-9; quiz 870. [doi: [10.1586/ern.12.63](https://doi.org/10.1586/ern.12.63)] [Medline: [22853793](https://pubmed.ncbi.nlm.nih.gov/22853793/)]

28. Palmqvist B, Carlbring P, Andersson G. Internet-delivered treatments with or without therapist input: does the therapist factor have implications for efficacy and cost? *Expert Rev Pharmacoecon Outcomes Res* 2007 Jun;7(3):291-297. [doi: [10.1586/14737167.7.3.291](https://doi.org/10.1586/14737167.7.3.291)] [Medline: [20528315](#)]
29. Sethi S, Campbell AJ, Ellis LA. The use of computerized self-help packages to treat adolescent depression and anxiety. *Journal of Technology in Human Services* 2010 Aug 31;28(3):144-160. [doi: [10.1080/15228835.2010.508317](https://doi.org/10.1080/15228835.2010.508317)]
30. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. *Aust N Z J Psychiatry* 2009 Jun;43(6):571-578. [doi: [10.1080/00048670902873722](https://doi.org/10.1080/00048670902873722)] [Medline: [19440890](#)]
31. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. *Psychol Med* 2010 Dec;40(12):1943-1957. [doi: [10.1017/S0033291710000772](https://doi.org/10.1017/S0033291710000772)] [Medline: [20406528](#)]
32. Groves C, Griffiths K, Christensen H. *Working out MoodGYM: A User's Guide*. Canberra, Australia: Centre for Mental Health Research; 2003.
33. Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. *J Med Internet Res* 2013;15(1):e2 [FREE Full text] [doi: [10.2196/jmir.2240](https://doi.org/10.2196/jmir.2240)] [Medline: [23302475](#)]
34. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the Internet: randomised controlled trial. *BMJ* 2004 Jan 31;328(7434):265 [FREE Full text] [doi: [10.1136/bmj.37945.566632.EE](https://doi.org/10.1136/bmj.37945.566632.EE)] [Medline: [14742346](#)]
35. Caelear AL, Christensen H, Mackinnon A, Griffiths KM, O'Kearney R. The YouthMood Project: a cluster randomized controlled trial of an online cognitive behavioral program with adolescents. *J Consult Clin Psychol* 2009 Dec;77(6):1021-1032. [doi: [10.1037/a0017391](https://doi.org/10.1037/a0017391)] [Medline: [19968379](#)]
36. Mackinnon A, Griffiths KM, Christensen H. Comparative randomised trial of online cognitive-behavioural therapy and an information website for depression: 12-month outcomes. *Br J Psychiatry* 2008 Feb;192(2):130-134 [FREE Full text] [doi: [10.1192/bjp.bp.106.032078](https://doi.org/10.1192/bjp.bp.106.032078)] [Medline: [18245031](#)]
37. Hickie IB, Davenport TA, Luscombe GM, Moore M, Griffiths KM, Christensen H. Practitioner-supported delivery of Internet-based cognitive behaviour therapy: evaluation of the feasibility of conducting a cluster randomised trial. *Med J Aust* 2010 Jun 7;192(11 Suppl):S31-S35. [Medline: [20528705](#)]
38. Wang PS, Berglund P, Kessler RC. Recent care of common mental disorders in the United States : prevalence and conformance with evidence-based recommendations. *J Gen Intern Med* 2000 May;15(5):284-292 [FREE Full text] [Medline: [10840263](#)]
39. Jorm AF. Mental health literacy. Public knowledge and beliefs about mental disorders. *Br J Psychiatry* 2000 Nov;177:396-401 [FREE Full text] [Medline: [11059991](#)]
40. Bebbington P, Meltzer H, Brugha T, Farrell M, Jenkins R, Ceresa C, et al. Unequal access and unmet need: neurotic disorders and the use of primary care services. *Int Rev Psychiatry* 2003;15(1-2):115-122. [doi: [10.1080/0954026021000046029](https://doi.org/10.1080/0954026021000046029)] [Medline: [12745318](#)]
41. Helsedirektoratet. Nasjonale retningslinjer for diagnostisering og behandling av voksne med depresjon i primær- og spesialisthelsetjenesten. 2009. URL: <http://www.helsedirektoratet.no/publikasjoner/nasjonale-retningslinjer-for-diagnostisering-og-behandling-av-voksne-med-depresjon-i-primær--og-spesialisthelsetjenesten/Sider/default.aspx> [accessed 2013-05-08] [WebCite Cache ID 6GStEoITt]
42. National Institute for Health and Clinical Excellence. Depression in adults: the treatment and management of depression in adults. 2009. URL: <http://www.nice.org.uk/nicemedia/live/12329/45888/45888.pdf> [accessed 2013-05-08] [WebCite Cache ID 6GStYHtL]
43. Angermeyer MC, Matschinger H. Public attitude towards psychiatric treatment. *Acta Psychiatr Scand* 1996 Nov;94(5):326-336. [Medline: [9124079](#)]
44. Priest RG, Vize C, Roberts A, Roberts M, Tylee A. Lay people's attitudes to treatment of depression: results of opinion poll for Defeat Depression Campaign just before its launch. *BMJ* 1996 Oct 5;313(7061):858-859 [FREE Full text] [Medline: [8870574](#)]
45. van Schaik DJ, Klijn AF, van Hout HP, van Marwijk HW, Beekman AT, de Haan M, et al. Patients' preferences in the treatment of depressive disorder in primary care. *Gen Hosp Psychiatry* 2004;26(3):184-189. [doi: [10.1016/j.genhosppsy.2003.12.001](https://doi.org/10.1016/j.genhosppsy.2003.12.001)] [Medline: [15121346](#)]
46. King M, Davidson O, Taylor F, Haines A, Sharp D, Turner R. Effectiveness of teaching general practitioners skills in brief cognitive behaviour therapy to treat patients with depression: randomised controlled trial. *BMJ* 2002 Apr 20;324(7343):947-950 [FREE Full text] [Medline: [11964340](#)]
47. Davidsen A. Experiences of carrying out talking therapy in general practice: a qualitative interview study. *Patient Educ Couns* 2008 Aug;72(2):268-275. [doi: [10.1016/j.pec.2008.03.020](https://doi.org/10.1016/j.pec.2008.03.020)] [Medline: [18472244](#)]
48. Backenstrass M, Joest K, Rosemann T, Szecsenyi J. The care of patients with subthreshold depression in primary care: is it all that bad? A qualitative study on the views of general practitioners and patients. *BMC Health Serv Res* 2007;7:190 [FREE Full text] [doi: [10.1186/1472-6963-7-190](https://doi.org/10.1186/1472-6963-7-190)] [Medline: [18031573](#)]
49. Wiebe E, Greiver M. Using cognitive behavioural therapy in practice: qualitative study of family physicians' experiences. *Can Fam Physician* 2005 Jul;51:992-993 [FREE Full text] [Medline: [16926952](#)]

50. Telford R, Hutchinson A, Jones R, Rix S, Howe A. Obstacles to effective treatment of depression: a general practice perspective. *Fam Pract* 2002 Feb;19(1):45-52 [FREE Full text] [Medline: [11818349](#)]
51. Mykletun A, Knudsen AK, Tangen T, Overland S. General practitioners' opinions on how to improve treatment of mental disorders in primary health care. Interviews with one hundred Norwegian general practitioners. *BMC Health Serv Res* 2010;10:35 [FREE Full text] [doi: [10.1186/1472-6963-10-35](#)] [Medline: [20144205](#)]
52. Blashki G, Parsons J, Morgan H, Hickie IB, Davenport TA. Providing psychological treatments in general practice. Rationale and practicalities. *Aust Fam Physician* 2003 Aug;32(8):625-630. [Medline: [12973871](#)]
53. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011;13(4):e126 [FREE Full text] [doi: [10.2196/jmir.1923](#)] [Medline: [22209829](#)]
54. 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](#)] [Medline: [19799804](#)]
55. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](#)] [Medline: [18812006](#)]
56. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio, TX: Psychological Corp; 1996.
57. Arnau RC, Meagher MW, Norris MP, Bramson R. Psychometric evaluation of the Beck Depression Inventory-II with primary care medical patients. *Health Psychol* 2001 Mar;20(2):112-119. [Medline: [11315728](#)]
58. Dozois DJA, Dobson KS, Ahnberg JL. A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment* 1998;10(2):83-89. [doi: [10.1037/1040-3590.10.2.83](#)]
59. Carlbring P, Brunt S, Bohman S, Austin D, Richards J, Öst LG, et al. Internet vs. paper and pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior* 2007 May;23(3):1421-1434. [doi: [10.1016/j.chb.2005.05.002](#)]
60. Holländare F, Andersson G, Engström I. A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *J Med Internet Res* 2010;12(5):e49 [FREE Full text] [doi: [10.2196/jmir.1392](#)] [Medline: [21169165](#)]
61. Naus MJ, Philipp LM, Samsi M. From paper to pixels: A comparison of paper and computer formats in psychological assessment. *Computers in Human Behavior* 2009 Jan;25(1):1-7. [doi: [10.1016/j.chb.2008.05.012](#)]
62. Schulenberg SE, Yutzenka BA. Equivalence of computerized and conventional versions of the Beck Depression Inventory-II (BDI-II). *Curr Psychol* 2001 Sep;20(3):216-230. [doi: [10.1007/s12144-001-1008-1](#)]
63. Beck AT, Steer RA. Beck Anxiety Inventory: Manual. San Antonio, TX: The Psychological Corporation; 1993.
64. Fydrich T, Dowdall D, Chambless DL. Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders* 1992 Jan;6(1):55-61. [doi: [10.1016/0887-6185\(92\)90026-4](#)]
65. Steer RA, Ranieri WF, Beck AT, Clark DA. Further evidence for the validity of the Beck Anxiety Inventory with psychiatric outpatients. *Journal of Anxiety Disorders* 1993 Jul;7(3):195-205. [doi: [10.1016/0887-6185\(93\)90002-3](#)]
66. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988 Dec;56(6):893-897. [Medline: [3204199](#)]
67. Hedman E, Ljótsson B, Rück C, Furmark T, Carlbring P, Lindfors N, et al. Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. *Computers in Human Behavior* 2010 Jul;26(4):736-740. [doi: [10.1016/j.chb.2010.01.010](#)]
68. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [Medline: [6880820](#)]
69. 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](#)]
70. Andersson G, Kaldo-Sandström V, Ström L, Strömgren T. Internet administration of the Hospital Anxiety and Depression Scale in a sample of tinnitus patients. *J Psychosom Res* 2003 Sep;55(3):259-262. [Medline: [12932800](#)]
71. McCue P, Buchanan T, Martin CR. Screening for psychological distress using Internet administration of the Hospital Anxiety and Depression Scale (HADS) in individuals with chronic fatigue syndrome. *Br J Clin Psychol* 2006 Nov;45(Pt 4):483-498. [doi: [10.1348/014466505X82379](#)] [Medline: [17076959](#)]
72. EuroQol Group. About EQ-5D. 2013. URL: <http://www.euroqol.org/about-eq-5d.html> [accessed 2013-05-08] [WebCite Cache ID [6GSukzvVz](#)]
73. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997 Nov;35(11):1095-1108. [Medline: [9366889](#)]
74. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health patients. *Health Econ* 2006 Nov;15(11):1229-1236. [doi: [10.1002/hec.1125](#)] [Medline: [16625671](#)]
75. Shervin N, Dorrwachter J, Bragdon CR, Shervin D, Zurakowski D, Malchau H. Comparison of paper and computer-based questionnaire modes for measuring health outcomes in patients undergoing total hip arthroplasty. *J Bone Joint Surg Am* 2011 Feb 2;93(3):285-293. [doi: [10.2106/JBJS.I.01104](#)] [Medline: [21266642](#)]
76. Bushnell DM, Reilly MC, Galani C, Martin ML, Ricci JF, Patrick DL, et al. Validation of electronic data capture of the Irritable Bowel Syndrome--Quality of Life Measure, the Work Productivity and Activity Impairment Questionnaire for

- Irritable Bowel Syndrome and the EuroQol. *Value Health* 2006;9(2):98-105. [doi: [10.1111/j.1524-4733.2006.00087.x](https://doi.org/10.1111/j.1524-4733.2006.00087.x)] [Medline: [16626413](https://pubmed.ncbi.nlm.nih.gov/16626413/)]
77. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess* 1985 Feb;49(1):71-75. [doi: [10.1207/s15327752jpa4901_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
78. Pavot W, Diener E. Review of the Satisfaction With Life Scale. *Psychological Assessment* 1993;5(2):164-172. [doi: [10.1037/1040-3590.5.2.164](https://doi.org/10.1037/1040-3590.5.2.164)]
79. Clench-Aas J, Nes RB, Dalgard OS, Aarø LE. Dimensionality and measurement invariance in the Satisfaction with Life Scale in Norway. *Qual Life Res* 2011 Oct;20(8):1307-1317 [FREE Full text] [doi: [10.1007/s11136-011-9859-x](https://doi.org/10.1007/s11136-011-9859-x)] [Medline: [21308414](https://pubmed.ncbi.nlm.nih.gov/21308414/)]
80. Glaesmer H, Grande G, Braehler E, Roth M. The German Version of the Satisfaction With Life Scale (SWLS): psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment* 2011 May 3;27(2):127-132. [doi: [10.1027/1015-5759/a000058](https://doi.org/10.1027/1015-5759/a000058)]
81. Meyerson P, Tryon WW. Validating Internet research: a test of the psychometric equivalence of Internet and in-person samples. *Behav Res Methods Instrum Comput* 2003 Nov;35(4):614-620. [Medline: [14748506](https://pubmed.ncbi.nlm.nih.gov/14748506/)]
82. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33;quiz 34. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
83. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction* 1993 Jun;88(6):791-804. [Medline: [8329970](https://pubmed.ncbi.nlm.nih.gov/8329970/)]
84. Reinert DF, Allen JP. The alcohol use disorders identification test: an update of research findings. *Alcohol Clin Exp Res* 2007 Feb;31(2):185-199. [doi: [10.1111/j.1530-0277.2006.00295.x](https://doi.org/10.1111/j.1530-0277.2006.00295.x)] [Medline: [17250609](https://pubmed.ncbi.nlm.nih.gov/17250609/)]
85. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care 2nd edition. 2001. URL: http://whqlibdoc.who.int/hq/2001/who_msd_msb_01.6a.pdf [accessed 2013-05-10] [WebCite Cache ID 6GVRHvzpE]
86. Berman AH, Bergman H, Palmstierna T, Schlyter F. Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *Eur Addict Res* 2005;11(1):22-31. [doi: [10.1159/000081413](https://doi.org/10.1159/000081413)] [Medline: [15608468](https://pubmed.ncbi.nlm.nih.gov/15608468/)]
87. Cavanagh K, Shapiro DA, Van Den Berg S, Swain S, Barkham M, Proudfoot J. The acceptability of computer-aided cognitive behavioural therapy: a pragmatic study. *Cogn Behav Ther* 2009;38(4):235-246. [doi: [10.1080/16506070802561256](https://doi.org/10.1080/16506070802561256)] [Medline: [19306147](https://pubmed.ncbi.nlm.nih.gov/19306147/)]
88. Garratt A, Bjørngaard JH, Dahle KA, Bjertnaes OA, Saunes IS, Ruud T. The Psychiatric Out-Patient Experiences Questionnaire (POPEQ): data quality, reliability and validity in patients attending 90 Norwegian clinics. *Nord J Psychiatry* 2006;60(2):89-96. [doi: [10.1080/08039480600583464](https://doi.org/10.1080/08039480600583464)] [Medline: [16635926](https://pubmed.ncbi.nlm.nih.gov/16635926/)]
89. Nasjonalt kunnskapssenter for helsetjenesten. PasOpp: hva er erfaringene dine som bruker av poliklinikk i psykiatrisk helsevern?. 2007. URL: <http://www.kunnskapssenteret.no/Sp%C3%B8rreskjemabanken/Spesialisthelsetjeneste++Psykisk+helsevern/Voksne%2C+psykisk+helsevern%2C+poliklinikk.9716.cms> [accessed 2013-05-10] [WebCite Cache ID 6GVRtgBlw]
90. Australian National University, Centre for Mental Health Research. The MoodGYM training program: mark III. 2008. URL: <https://moodgym.anu.edu.au/welcome> [accessed 2013-05-10] [WebCite Cache ID 6GVUeWUTc]
91. Christensen H, Griffiths KM, Mackinnon AJ, Brittliffe K. Online randomized controlled trial of brief and full cognitive behaviour therapy for depression. *Psychol Med* 2006 Dec;36(12):1737-1746. [doi: [10.1017/S0033291706008695](https://doi.org/10.1017/S0033291706008695)] [Medline: [16938144](https://pubmed.ncbi.nlm.nih.gov/16938144/)]
92. Twisk JWR. Applied multilevel analysis: a practical guide. Cambridge, UK: Cambridge University Press; 2006.
93. Silfvernagel K, Carlbring P, Kabo J, Edström S, Eriksson J, Månson L, et al. Individually tailored Internet-based treatment for young adults and adults with panic attacks: randomized controlled trial. *J Med Internet Res* 2012;14(3):e65 [FREE Full text] [doi: [10.2196/jmir.1853](https://doi.org/10.2196/jmir.1853)] [Medline: [22732098](https://pubmed.ncbi.nlm.nih.gov/22732098/)]
94. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods* 2002 Mar;7(1):105-125. [Medline: [11928886](https://pubmed.ncbi.nlm.nih.gov/11928886/)]
95. Thorsell J, Finnes A, Dahl J, Lundgren T, Gybrant M, Gordh T, et al. A comparative study of 2 manual-based self-help interventions, acceptance and commitment therapy and applied relaxation, for persons with chronic pain. *Clin J Pain* 2011 Oct;27(8):716-723. [doi: [10.1097/AJP.0b013e318219a933](https://doi.org/10.1097/AJP.0b013e318219a933)] [Medline: [21540740](https://pubmed.ncbi.nlm.nih.gov/21540740/)]
96. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: L Erlbaum Associates; 1988.
97. Seggar LB, Lambert MJ, Hansen NB. Assessing clinical significance: Application to the Beck Depression Inventory. *Behavior Therapy* 2002 Mar;33(2):253-269. [doi: [10.1016/s0005-7894\(02\)80028-4](https://doi.org/10.1016/s0005-7894(02)80028-4)]
98. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991 Feb;59(1):12-19. [Medline: [2002127](https://pubmed.ncbi.nlm.nih.gov/2002127/)]

99. Choi I, Zou J, Titov N, Dear BF, Li S, Johnston L, et al. Culturally attuned Internet treatment for depression amongst Chinese Australians: a randomised controlled trial. *J Affect Disord* 2012 Feb;136(3):459-468. [doi: [10.1016/j.jad.2011.11.003](https://doi.org/10.1016/j.jad.2011.11.003)] [Medline: [22177742](https://pubmed.ncbi.nlm.nih.gov/22177742/)]
100. Vermark K, Lenndin J, Bjärehed J, Carlsson M, Karlsson J, Oberg J, et al. Internet administered guided self-help versus individualized e-mail therapy: A randomized trial of two versions of CBT for major depression. *Behav Res Ther* 2010 May;48(5):368-376. [doi: [10.1016/j.brat.2010.01.005](https://doi.org/10.1016/j.brat.2010.01.005)] [Medline: [20152960](https://pubmed.ncbi.nlm.nih.gov/20152960/)]
101. Barrett JE, Williams JW, Oxman TE, Frank E, Katon W, Sullivan M, et al. Treatment of dysthymia and minor depression in primary care: a randomized trial in patients aged 18 to 59 years. *J Fam Pract* 2001 May;50(5):405-412. [Medline: [11350703](https://pubmed.ncbi.nlm.nih.gov/11350703/)]
102. Hermens ML, van Hout HP, Terluin B, van der Windt DA, Beekman AT, van Dyck R, et al. The prognosis of minor depression in the general population: a systematic review. *Gen Hosp Psychiatry* 2004;26(6):453-462. [doi: [10.1016/j.genhosppsy.2004.08.006](https://doi.org/10.1016/j.genhosppsy.2004.08.006)] [Medline: [15567211](https://pubmed.ncbi.nlm.nih.gov/15567211/)]
103. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. *J Med Internet Res* 2002;4(1):e3 [FREE Full text] [doi: [10.2196/jmir.4.1.e3](https://doi.org/10.2196/jmir.4.1.e3)] [Medline: [11956035](https://pubmed.ncbi.nlm.nih.gov/11956035/)]
104. Goldberg D, Bridges K, Duncan-Jones P, Grayson D. Detecting anxiety and depression in general medical settings. *BMJ* 1988 Oct 8;297(6653):897-899 [FREE Full text] [Medline: [3140969](https://pubmed.ncbi.nlm.nih.gov/3140969/)]
105. Bados A, Balaguer G, Saldaña C. The efficacy of cognitive-behavioral therapy and the problem of drop-out. *J Clin Psychol* 2007 Jun;63(6):585-592. [doi: [10.1002/jclp.20368](https://doi.org/10.1002/jclp.20368)] [Medline: [17457848](https://pubmed.ncbi.nlm.nih.gov/17457848/)]
106. Watkins E, Williams R. The efficacy of cognitive behavioral therapy. In: Checkley S, editor. *The Management of Depression*. Oxford, UK: Blackwell Science; 1998:165-188.
107. Wierzbicki M, Pekarik G. A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice* 1993;24(2):190-195. [doi: [10.1037/0735-7028.24.2.190](https://doi.org/10.1037/0735-7028.24.2.190)]
108. Lamers F, van Oppen P, Comijs HC, Smit JH, Spinhoven P, van Balkom AJ, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry* 2011 Mar;72(3):341-348. [doi: [10.4088/JCP.10m06176blu](https://doi.org/10.4088/JCP.10m06176blu)] [Medline: [21294994](https://pubmed.ncbi.nlm.nih.gov/21294994/)]
109. Jacobi F, Wittchen HU, Holting C, Höfler M, Pfister H, Müller N, et al. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med* 2004 May;34(4):597-611. [doi: [10.1017/S0033291703001399](https://doi.org/10.1017/S0033291703001399)] [Medline: [15099415](https://pubmed.ncbi.nlm.nih.gov/15099415/)]
110. van Boeijen CA, van Oppen P, van Balkom AJ, Visser S, Kempe PT, Blankenstein N, et al. Treatment of anxiety disorders in primary care practice: a randomised controlled trial. *Br J Gen Pract* 2005 Oct;55(519):763-769 [FREE Full text] [Medline: [16212851](https://pubmed.ncbi.nlm.nih.gov/16212851/)]
111. Longwell BT, Truax P. The differential effects of weekly, monthly, and bimonthly administrations of the beck Depression Inventory-II: Psychometric properties and clinical implications. *Behavior Therapy* 2005 Jun;36(3):265-275 [FREE Full text] [doi: [10.1016/S0005-7894\(05\)80075-9](https://doi.org/10.1016/S0005-7894(05)80075-9)]
112. Hatzenbuehler LC, Parpal M, Matthews L. Classifying college students as depressed or nondepressed using the Beck Depression Inventory: an empirical analysis. *J Consult Clin Psychol* 1983 Jun;51(3):360-366. [Medline: [6863694](https://pubmed.ncbi.nlm.nih.gov/6863694/)]
113. Haaga DA. Introduction to the special section on stepped care models in psychotherapy. *J Consult Clin Psychol* 2000 Aug;68(4):547-548. [Medline: [10965628](https://pubmed.ncbi.nlm.nih.gov/10965628/)]
114. Farrand P, Confue P, Byng R, Shaw S. Guided self-help supported by paraprofessional mental health workers: an uncontrolled before--after cohort study. *Health Soc Care Community* 2009 Feb;17(1):9-17. [doi: [10.1111/j.1365-2524.2008.00792.x](https://doi.org/10.1111/j.1365-2524.2008.00792.x)] [Medline: [18564197](https://pubmed.ncbi.nlm.nih.gov/18564197/)]
115. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10939 [FREE Full text] [doi: [10.1371/journal.pone.0010939](https://doi.org/10.1371/journal.pone.0010939)] [Medline: [20544030](https://pubmed.ncbi.nlm.nih.gov/20544030/)]

Abbreviations

- AUDIT:** The Alcohol Use Disorder Identification Test
- ANOVA:** analysis of variance
- ANCOVA:** analysis of covariance
- ANU:** Australian National University
- BAI:** Beck Anxiety Inventory
- BDI-II:** Beck Depression Inventory, second edition
- CBT:** cognitive behavioral therapy
- DSM-IV:** Diagnostic and Statistical Manual of Mental Disorders, 4th edition
- DUDIT:** The Drug Use Disorder Identification Test
- EQ-5D:** EuroQol Group 5-Dimension Self-Report Questionnaire
- EQ Index:** index score of the EQ-5D

EQ VAS: visual analog scale of the EQ-5D
GP: general practitioner
HADS: Hospital Anxiety and Depression Scale
ICD-10: International Classification of Diseases, 10th edition
ITT: intention-to-treat
SWLS: Satisfaction with Life Scale

Edited by G Eysenbach; submitted 19.05.13; peer-reviewed by T Berger, T Donker; comments to author 10.06.13; revised version received 03.07.13; accepted 12.07.13; published 05.08.13.

Please cite as:

*Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, Kolstrup N
The Clinical Effectiveness of Web-Based Cognitive Behavioral Therapy With Face-to-Face Therapist Support for Depressed Primary Care Patients: Randomized Controlled Trial
J Med Internet Res 2013;15(8):e153
URL: <http://www.jmir.org/2013/8/e153/>
doi: [10.2196/jmir.2714](https://doi.org/10.2196/jmir.2714)
PMID: [23916965](https://pubmed.ncbi.nlm.nih.gov/23916965/)*

©Ragnhild Sørensen Høifødt, Kjersti R Lillevoll, Kathleen M Griffiths, Tom Wilsgaard, Martin Eisemann, Knut Waterloo, Nils Kolstrup. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 05.08.2013. 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

Staging Dementia From Symptom Profiles on a Care Partner Website

Kenneth Rockwood^{1,2}, MD, FRCPC; Matthew Richard², BSc; Chris Leibman³, MS, PharmD; Lisa Mucha⁴, PhD; Arnold Mitnitski^{1,2}, PhD

¹Dalhousie University, Department of Medicine, Dalhousie University, Halifax, NS, Canada

²DGI Clinical Inc, Halifax, NS, Canada

³Janssen Alzheimer Immunotherapy Research and Development, LLC, South San Francisco, CA, United States

⁴Pfizer Inc, Collegeville, PA, United States

Corresponding Author:

Kenneth Rockwood, MD, FRCPC

DGI Clinical Inc

1344 Summer Street

Suite 208

Halifax, NS, B3H 0A8

Canada

Phone: 1 902 421 5710 ext 5

Fax: 1 902 421 2733

Email: krockwood@dgiclinical.com

Abstract

Background: The World Wide Web allows access to patient/care partner perspectives on the lived experience of dementia. We were interested in how symptoms that care partners target for tracking relate to dementia stage, and whether dementia could be staged using only these online profiles of targeted symptoms.

Objectives: To use clinical data where the dementia stage is known to develop a model that classifies an individual's stage of dementia based on their symptom profile and to apply this model to classify dementia stages for subjects from a Web-based dataset.

Methods: An Artificial Neural Network (ANN) was used to identify the relationships between the dementia stages and individualized profiles of people with dementia obtained from the 60-item SymptomGuide (SG). The clinic-based training dataset (n=320), with known dementia stages, was used to create an ANN model for classifying stages in Web-based users (n=1930).

Results: The ANN model was trained in 66% of the 320 Memory Clinic patients, with the remaining 34% used to test its accuracy in classification. Training and testing staging distributions were not significantly different. In the 1930 Web-based profiles, 309 people (16%) were classified as having mild cognitive impairment, 36% as mild dementia, 29% as moderate, and 19% as severe. In both the clinical and Web-based symptom profiles, most symptoms became more common as the stage of dementia worsened (eg, mean 5.6 SD 5.9 symptoms in the MCI group versus 11.9 SD 11.3 in the severe). Overall, Web profiles recorded more symptoms (mean 7.1 SD 8.0) than did clinic ones (mean 5.5 SD 1.8). Even so, symptom profiles were relatively similar between the Web-based and clinical datasets.

Conclusion: Symptoms targeted for online tracking by care partners of people with dementia can be used to stage dementia. Even so, caution is needed to assure the validity of data collected online as the current staging algorithm should be seen as an initial step.

(*J Med Internet Res* 2013;15(8):e145) doi:[10.2196/jmir.2461](https://doi.org/10.2196/jmir.2461)

KEYWORDS

dementia; staging; online survey; symptoms; Cognitive Impairment Not Dementia; Mild Cognitive Impairment; validation; Artificial Neural Networks; World Wide Web

Introduction

The World Wide Web offers new opportunities for understanding disease from a patient's standpoint, and crucially in dementia, from the standpoint of their caregivers [1]. On the Web, detailed information can be collected from survey data [2,3] or extracted from online programs offered to caregivers [4-8]. In any examination of patient/caregiver perspectives on the lived experience of dementia, understanding the stage of dementia being discussed is crucial. Unfortunately, how best to stage dementia using caregiver reports on the Internet is not clear. Earlier, we have shown that a structured questionnaire based on the Dependence Scale [9] designed to grade increasing degrees of dependence showed good construct validity as a staging measure [2]. While structured questionnaires can be employed, users can see them as intrusive and unrewarding, especially if their completion seems to require undue effort. Even so, it is very important for caregivers to have some sense about dementia stage since many disease manifestations are stage dependent (eg, wandering is a later stage symptom); whereas in other cases, a symptom appearing "out of order" (in relation to untreated Alzheimer's disease) would have diagnostic value (eg, hallucinations occurring very early in the dementia course would suggest Lewy Body disease).

The SymptomGuide for dementia is a Web-based tool aimed particularly at the caregivers of persons with dementia. This tool allows a caregiver to track the health status and symptoms of the person they are giving care to. In addition, using its corresponding online symptom library, caregivers can learn more about common manifestations of dementia [10]. We were interested in using this information as a means of staging dementia. Of the many instruments commonly used clinically to stage dementia, none relies only on symptoms. Constructing algorithms to stage dementia from symptoms alone is fundamentally challenging using a priori rules since many common symptoms can occur at different stages of dementia. In general, the complex relationships between clearly associated symptoms and dementia stages are difficult to discern using conventional classical statistical methods. By contrast, artificial intelligence systems can approximate the complex nonlinear relationships between these variables including outcomes [11]. Artificial Neural Network (ANN) machine-learning techniques could be particularly beneficial in discovering patterns and how they change with disease progression. ANNs have been applied in the discrimination of mild cognitive impairment (MCI) from Alzheimer's disease [12] and to identify risk factors on the conversion of amnesic mild cognitive impairment [13], as well as in analysis of neuroimaging data [14,15]. ANNs have also been applied to associating individual characteristics with outcomes [16]. ANNs have been compared with conventional statistical approaches [17] and in particular in Alzheimer's disease research [18], suggesting the usefulness of this approach.

Our overall objective was to develop an application of a machine-learning ANN algorithm to stage dementia using Web-based, individualized symptom profiles. In particular, we aimed to (1) develop and validate a symptom-based staging system using memory clinic data, where staging can be verified, (2) apply this to the Web-based symptom data, and (3) explore

differences in online and memory clinic symptom targeting that might influence staging and its interpretation.

Methods

Setting

The data came from the SymptomGuide (SG) website (see [Multimedia Appendix 1](#)). Caregivers can visit this site to learn about symptoms exhibited by the person they are caring for. Individuals with dementia can also input data; although, only 1% of this site's users report having dementia themselves. Crucially for these analyses, caregivers target the symptoms most relevant to them in order to track the course of the disease and/or the effects of treatment [11] (people being profiled on the Web may have many more symptoms than those being targeted for tracking; these profiles need to be interpreted as statements about the most troubling symptoms and not as symptom inventories). The online symptom library ([Multimedia Appendix 2](#)) describes 60 symptoms, each detailed using about a dozen plain-language descriptors [19]. The library defines and describes each symptom and includes information about the typical stage of dementia when that symptom occurs. Subcategories in the library provide other relevant information, accessed by clicking on tabs visible for each symptom (eg, a tab entitled "Doctor's Diary" provides standard advice from a physician about the typical challenges and course related to that problem). One heavily trafficked subsection describes common management strategies that can be employed in relation to each symptom. For users who build symptom profiles, learning about and tracking symptoms is their chief interaction with the site, so that building a staging algorithm from the patient profiles does not require additional effort by care partners.

Between its launch in 2007 and March 2012, 6129 online users have built symptom profiles, of whom 1930 have also created complete personal profiles, which consist of data about demographics, medications, symptoms, and symptom progression. These data were considered the Web-based data for this study.

Measures

In addition to Web-based users, the SG is used in a tertiary care Memory Clinic in Halifax, Nova Scotia. Each of the 323 clinic-based SG users whose data are considered here underwent standard assessments, which included staging based on the Global Deterioration Scale (GDS) score for dementia [20] (determined during the clinical interview by the examining physician as the mean value of the first 5 axes of the Brief Cognitive Rating Scale [21]). The GDS was scored as 3=cognitive impairment not dementia (CIND)/mild cognitive impairment, 4=mild dementia, 5=moderate dementia, and 6=severe dementia. A Mini-Mental State Examination (MMSE, usually administered by a clinic nurse) was also scored. The MMSE is a screening cognitive test, scored from 0-30, with a higher score indicating better performance. The GDS rater was an experienced clinician scientist (the first author, KR) who was not blind to the MMSE. Earlier work had suggested that a detailed questionnaire for staging was not of interest to most users. Here we substituted a brief questionnaire that asked if the person being profiled has been diagnosed with dementia

and then described each stage in a single sentence. Users were asked which sentence best described the person being profiled.

Analysis

A symptom-derived staging algorithm is proposed to recognize four major groupings of cognitive impairment: MCI, mild, moderate, and severe dementia. The algorithm was developed (Objective 1) from profiles in the Memory Clinic (where staging is known) and applied to the Web dataset (where staging usually is not contained in the database).

To develop the algorithm, we used an ANN, which consists of processing units that are called “neurons” because of certain similarities with human neurons that respond to input stimuli in a nonlinear fashion. The algorithm recursively analyzes their ability to predict an individual’s dementia stage, given the information about that individual’s symptoms and staging. The ANN was applied to a random selection of 66% (211/320) of the Memory Clinic database profiles to train the ANN. The remaining 34% of the clinic sample was used to assess the accuracy of the ANN model. The input variables for the ANN model were the presence or absence of SG symptoms and the person’s stage. Of the 60 symptoms presented in the SG, we used the 34 that had been used in at least 5% of the Web dataset and at least 5 times in the clinical dataset. Parameters of the ANN model included 4 nodes, an over fit penalty of 0.15, a 0.00001 convergence criteria, and 5 tours of 500 iterations. The output variables were the predicted probabilities of the 4 dementia stages. These specific ANN model parameters were chosen to optimize the percent of correctly predicted stages in the test data (ie, the 34% of clinic records not used in the training set). To test the robustness of the model, each training session was repeated 30 times; stability was tested using the coefficient of variation, with a tolerance of 15% change in classification.

The ANN model found using the clinic data was cross-validated in that dataset by correlating it with the MMSE and presentation, box-plot diagrams (Objective 1). It was then used to predict the stage of dementia in the Web-based sample (N=1930), for whom no stage of dementia was known (Objective 2). To explore how symptom targeting online might differ from symptom targeting in the Memory Clinic (Objective 3), we first cross-tabulated symptom profiles by stage of dementia for both the known clinical dataset and the predicted Web-based dataset. Next, we compared the number of symptoms set in each sample, again by dementia stage. We also explored how commonly Analysis of Variance (Kruskal-Wallis ANOVA) was used to measure group mean differences in MMSE by stage. The Pearson chi-square test was used to test for differences between staging distributions, using a significance level of 0.01. Calculations and analysis were performed using R statistical software v2.14.2.

Ethics

The study was approved by the Research Ethics Committee of the Capital District Health Authority, Nova Scotia. Clinic participants signed informed consent. All respondents to the survey consented by checking their agreement to a terms and

conditions list, which included their consent to the use of anonymized data. No personal information was collected that could identify the survey participant. All responses are stored on a secure server.

Results

Memory Clinic

The Memory Clinic patients were of a similar age to Web users, but more of the latter were women, and fewer Web users lived with family members (Table 1).

The machine-learning algorithm developed in the Memory Clinic training dataset showed virtually the same dementia stage distribution when applied to the testing dataset as did the clinical dementia staging. By both staging assignments, most patients had mild dementia (55% and 61%, respectively), followed by CIND/MCI (23% and 20%), moderate dementia (14%, 12%), and severe dementia (8%, 7%). Each training session was repeated 30 times; the coefficient of variation never exceeded 12% for each stage. The final ANN model showed a misclassification rate of 3%. In the Memory Clinic dataset, the ANN staging algorithm was significantly related to the MMSE scores ($F_{3,26}=101.1$, $P<.001$) (Figure 1).

In the Web-based dataset, most people were in the mild stage (Figure 2).

Comparison of Symptom Profiling Between the Memory Clinic and the Online Datasets

Symptom profiles were relatively similar between the Web-based and clinical datasets (Figure 3). Three trends were evident. First, even the most common symptoms selected for tracking in the Web-based dataset occur in less than half the profiles. This is also true for all but two symptoms in the Memory Clinic dataset. Even so, in general, the symptom profile of the Web-based dataset showed slightly higher symptom occurrence rates when compared to the clinical dataset. Overall, people who used the website targeted more symptoms (mean 7.1 SD 8.0) than did people in the Memory Clinic (5.5 SD 1.8) ($t_{2101}=-7.69$, [Welch’s t test for unequal sample sizes and unequal variances], $P<.001$). This appears to arise as a consequence of the third trend, which is that symptom targeting rates increased as the dementia severity stage progressed into stages 5 and 6. Specifically, Web users whose profiles conformed to stage 3 had a mean 5.6 ± 4.9 symptoms vs people in Memory Clinic, (4.97 SD 1.94, $t_{285}=-1.73$, $P=.083$); those in stage 4 targeted 4.8 SD 5.6 symptoms vs clinical mean 5.7 SD 1.61 ($t_{854}=3.69$, $P<.001$); in stage 5, 7.8 SD 7.9 symptoms vs clinical mean 6.07 SD 2.22 ($t_{173}=-3.71$, $P<.001$); and in stage 6, 11.9 SD 11.3 vs clinical mean 5.67 SD 1.97 ($t_{17}=-8.72$, $P<.001$). Given that the Web-based dataset had more patients in the moderate (29%) and the severe (19%) stages than did the Memory Clinic dataset (14% moderate and 8% severe), this appears to account for the difference in the mean number of symptoms between the two groups (chi-square₃=71.3, $P<.001$).

Table 1. Demographic characteristics of Memory Clinic patients and Web users.

Demographics	Memory clinic data N=320	Web-based data N=1930
Mean age, years (SD)	72 (10.4)	74 (10.4)
% female	46	59
Living arrangements, %		
Alone	7	20
With spouse/family or friend	90	58
Care facility/ nursing home	3	22

Figure 1. The association between Mini-Mental State Examination scores and dementia stages in Memory Clinic patients (N=109) staged according to ANN staging algorithm.

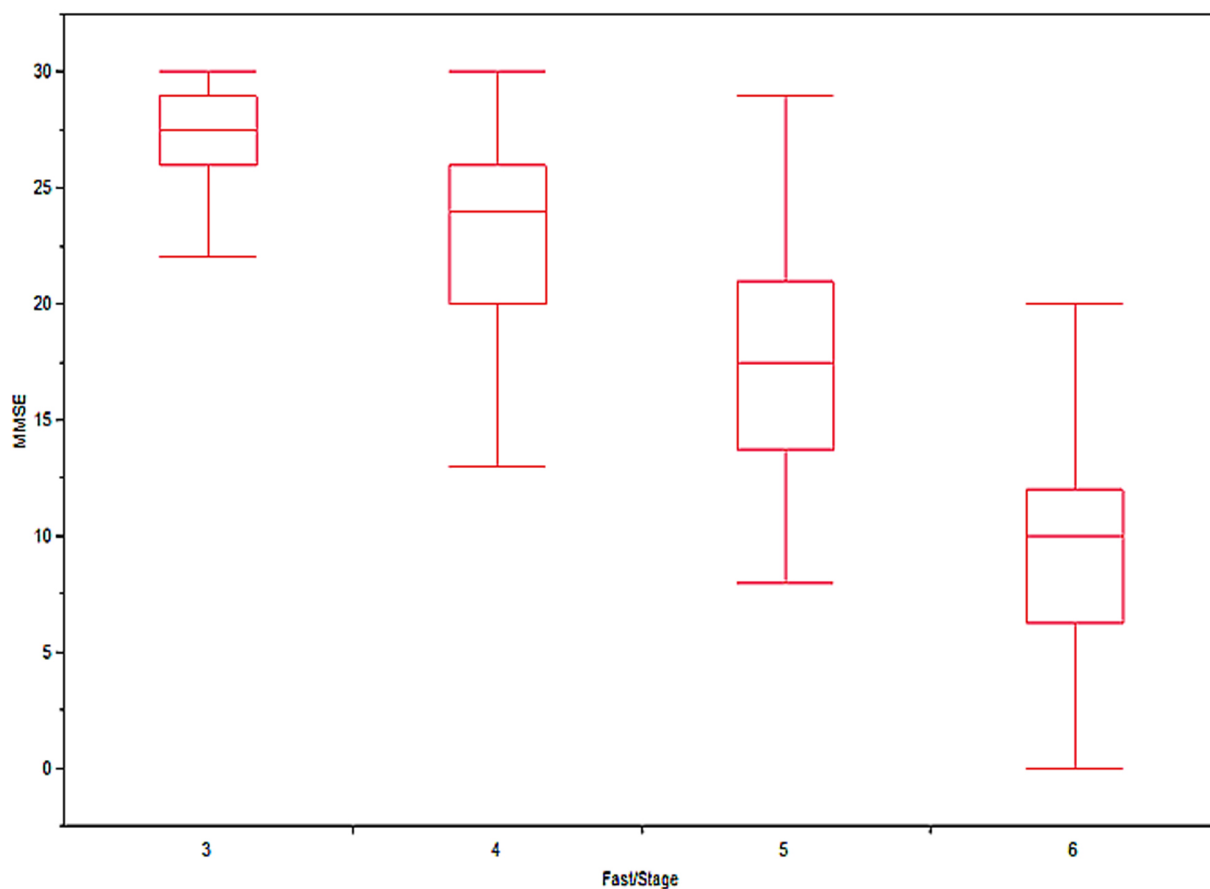


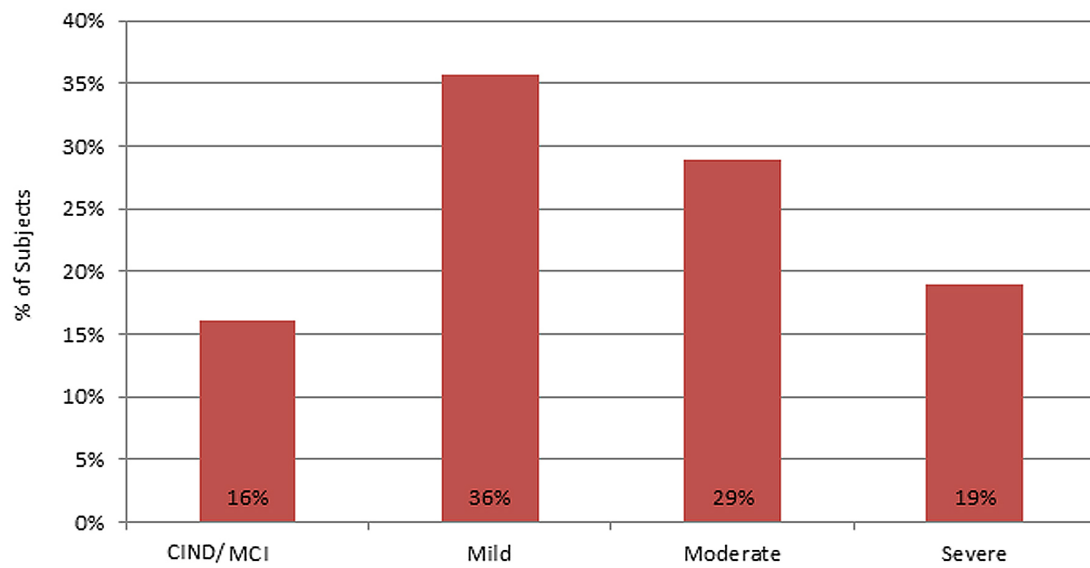
Figure 2. Distribution of Web-based users by clinical stages as classified by the ANN model staging algorithm (n=1930).

Figure 3. Comparison of the distribution of symptoms between the Memory Clinic and Web-users. Colour represents symptom frequencies: green (lower frequencies) to yellow (intermediate) to red (higher frequencies).

Symptoms	Stage Classification								
	Memory Clinic Dataset					Web-based Dataset			
	3	4	5	6		3	4	5	6
Incontinence	0%	2%	4%	17%		6%	12%	18%	29%
Inappropriate Language and Behaviour	0%	0%	0%	0%		9%	6%	10%	15%
Obsessive Behaviour	0%	1%	0%	0%		2%	4%	6%	17%
Repetitive Behaviour	0%	1%	2%	0%		4%	6%	8%	14%
Restlessness	0%	1%	0%	21%		2%	3%	7%	16%
Unsafe Actions	0%	2%	0%	0%		2%	6%	8%	15%
Aggression	1%	3%	4%	17%		12%	8%	12%	18%
Decline in Hobbies	1%	3%	0%	0%		4%	8%	9%	17%
Disorientation to Place	1%	2%	13%	4%		2%	4%	13%	11%
Impaired Memory for Names and Faces	1%	1%	2%	0%		4%	7%	22%	24%
Needs help with Dressing	1%	4%	9%	4%		4%	5%	13%	13%
Needs help with Shopping	1%	1%	0%	0%		5%	4%	9%	14%
Problems Operating Gadgets/Appliances	1%	5%	7%	4%		0%	7%	16%	16%
Low Mood	3%	1%	0%	0%		30%	8%	14%	23%
Wandering	3%	1%	7%	25%		3%	2%	7%	9%
Balance problems	4%	2%	7%	4%		13%	7%	17%	25%
Delusions and Paranoia	4%	8%	17%	25%		9%	9%	11%	18%
Hallucinations	4%	11%	20%	21%		2%	6%	9%	7%
Disorientation to Time	6%	7%	9%	13%		3%	9%	17%	13%
Problems with Household Chores	6%	12%	7%	8%		4%	7%	12%	19%
Problems with Insight	7%	6%	7%	8%		5%	4%	7%	7%
Problems with Driving	8%	11%	4%	0%		10%	6%	8%	18%
Problems Following Instructions	10%	6%	4%	4%		8%	7%	16%	20%
Poor Personal Care/Hygiene	11%	6%	28%	25%		20%	3%	19%	16%
Decreased Interest/Initiative	11%	11%	9%	8%		10%	11%	11%	18%
Problems with Telephone Use	13%	25%	28%	8%		9%	7%	15%	20%
Misplacing or Losing Objects	14%	23%	15%	4%		18%	18%	29%	31%
Past Memory Impairment	15%	13%	22%	21%		6%	6%	13%	20%
Sleep Disturbances	17%	15%	24%	29%		16%	16%	19%	37%
Impaired Judgment	18%	24%	41%	21%		7%	7%	10%	17%
Anxiety and Worry	20%	10%	11%	17%		15%	11%	9%	30%
Irritability/Frustration	24%	25%	33%	42%		25%	16%	26%	42%
Social Interaction/ Withdrawal	24%	21%	24%	25%		21%	9%	12%	29%
Repetitive Questions/ Stories	55%	67%	50%	25%		33%	28%	15%	34%
Memory of Recent Events	77%	80%	61%	46%		26%	25%	27%	27%

Discussion

Principal Findings

This paper used a machine-learning, symptom-based staging algorithm, developed from a Memory Clinic database of symptoms and stages, to define stages using only symptoms in a Web-based dataset. The Web-based dataset recorded symptoms targeted for treatment by care providers of people with dementia (and in a very few cases, by people with dementia themselves). Using an ANN model with 34 symptoms as inputs, consistent dementia stage classification was possible. The staging algorithm was trained and its accuracy tested using clinical data, where the stage of dementia for each patient was established according to the standard diagnostic. The misclassification rate in the Memory Clinic data was low (3%) and indicated a very good performance of the ANN model when validated internally (ie, using the Memory Clinic data). It was not possible to validate the model externally (in the Web-based data) in the same way as was done in the Memory Clinic because different criteria were used in both datasets to define the stages. In the Memory Clinic, staging was done by clinicians using the judgment-based GDS.

ANN models are known to be powerful machine-learning techniques which, when properly trained and tested, can give reliable predictions for unknown variables of interest. Until comparatively recently, their applicability in medical research was fairly limited due to the restriction of computer processing power and the lack of special training. Recent versions of major statistical packages now allow for user-friendly ANN analysis options (eg, JMP, MATLAB, Weka, R). Another major drawback in the application of ANNs is the difficulty in interpreting the results. This is because complex nonlinear relationships do not yield simple interpretations of the relationships between input variables (eg, symptoms) and response variables in a cause-effect manner. While such “black box” techniques show abundant applications in engineering, technical physics, and computer science, they are often less favorably received in the biomedical community, as they do not provide insight into relationships among variables. Even so, despite the desirable goal of understanding relationships among variables, the high dimensionality of problems like dementia (the dimensionality here is represented by many symptoms) makes such links nearly impossible. In situations like this, the application of ANNs presents a viable alternative for bypassing the immense complexity issue of our data and creating a model that, even though the relationship is unknown, can still reliably predict response variables by properly training and testing subsets of our data. Of note here, the simplest training (based only on associations within the symptom profile) was equally as informative as more complex algorithms, such as ones employing symptom severity, domain aggregation, and ratio of domain frequencies (eg, ratio of symptoms in the functional domain to symptoms in the behavioral domain).

The usual contrast to “bioinformatic” techniques such as ANNs, is to use “biostatistical” ones, such as factor analysis or its variant. The latter approaches, however, can sometimes ignore items that can be highly informative for individuals but are not

“statistically significant” at the group level. The additive value of effects which themselves can be negligible was recently illustrated in dementia epidemiology, in which a risk factor index made up of items that did not significantly predict dementia individually was more powerful than any single traditional risk factor in dementia prediction [22]. In short, in situations of high dimensionality, tradeoffs will be needed according to analytical intent. Here, the intent is to include the patient/caregiver perspective, using as much information as possible.

Our data should be interpreted with caution. The results reflect experience in using this emerging technology (in developing a model) rather than a claim to have developed a perfectly valid and accurate model. In short, the current model should be seen as an initial step. More specifically, despite similar symptom patterns in the Web-based data and the more controlled clinical environment data, one does not map exactly to the other. On the other hand, these differences may possibly reflect actual dementia severity differences between Web-based users and clinic patients. More studies should be done in order to better understand if this is the case. Likewise, for model stability, we used only symptoms that had been used at least 5 times in the training dataset, resulting in 26 of the SG’s 60 symptoms not being used in the staging algorithm. When the training dataset includes >1000 people, we plan to reassess the algorithm to evaluate its stability and the impact of less common symptoms on staging.

Further changes to the website now allow users to make their own staging assessment based on functional, behavioral, and cognitive symptom profiles, and this too will be re-evaluated periodically. In consequence, the current staging algorithm should be seen as an initial step. Even so, the initial results suggest caution in brief summary staging measures. In contrast to the more detailed Dependence Scale, which was cumbersome for many users, a very brief staging method did not improve uptake and made precision worse. A brief staging questionnaire, which described each stage in a single sentence, was completed by only 207 people, with weak ($r=0.32$) Spearman correlation with the ANN staging. Overall, the algorithm classified 37% of people into their observer-assessed category; this improved to 83% being classified within one level. Misclassification was normally distributed about 0. On the other hand, this lack of agreement may itself be informative. While it is the case that descriptions that are standard but brief enough to be completed by users may lack validity, the discrepancy between the staging algorithm and the brief questionnaire might in fact reflect the effect of treatment. The clinic-based profiles were weighted to patients prior to treatment (ie, at the time of initial diagnosis) but include many people who have been on treatment for months to years who are being reassessed. This would also be true of Web-based users. Given that currently used medications can alleviate some symptoms but do not cure or even halt progression, the stages detected by that algorithm might correspond less well to staging based on the untreated natural history, as was the case with the brief questionnaire. In addition, in each case we are mainly looking at the caregiver’s impression of the symptoms, especially if the dementia is more than mild. Very few profiles (<1%) appear to be completed by the patient

alone, and these are mostly weighted to those with MCI. The extent to which patients' insights into their own deficits might invalidate their own accounts (or even influence their caregivers) is not known.

Even so, it is inherent in a Web database that less strict quality control is possible compared to a clinical database. In consequence, the information needs to be interpreted with caution. However, because the Web database allows a less medicalized interaction for users, it may offer additional insights on the lived experience of dementia. Clinic-based datasets rarely assay symptoms beyond what exists in standardized scales and checklists, which typically do not include the same depth of information as here and typically record information on fewer people. This is an inherent trade-off, but the current experience suggests that the Web has great potential to provide useful information. In this regard, we were struck that, taking into account stage differences, the patterns were generally similar between clinics, where symptom choice is more influenced by interactions with health care professionals and completion of standard questionnaires, and online, where it appears that most people are doing this at home without such prompting.

Being able to stage dementia using Web profiles is useful in lessening the response burden of users. It also allows naturally occurring profiles to be used, enhancing the user's sense of contribution, instead of just completing questionnaires. More importantly, as more people with dementia are now being treated, many of the traditional staging algorithms need to be revisited. No current treatment is curative, so different combinations of mild, moderate, and severe staging items are

seen, especially in patients who have been on treatment for more than a year or two. As the database grows, it should be possible to explore these relationships better. Of note, the clinical dataset that trained the algorithm included symptoms for people both receiving and not receiving treatment, so it reflects this new reality.

We found it interesting that, compared with the Memory Clinic database, in the Web-based data, symptom targeting rates increased as the dementia severity stage increased. This may reflect that Web-based users have more severe problems compared to Memory Clinic patients. Alternately, the Memory Clinic patients may have these problems too, but they are not being targeted; this is a proposition that needs to be tested.

Conclusions

In general, robust classification of such a large sample of Web-based users allows for additional studies to be performed that reflect this perspective, including people who do not have access to memory clinic services. If further validated, it can provide a self-assessing staging classification that a caregiver can perform without additional training. Even so, lack of a means of verifying information is one reason that online data must be treated with caution. Finally, especially as disease-modifying drugs are developed that modify the course of dementia (and thereby its stages), it could lead to the creation of a more robust clinical staging methodology that considers symptom profile composition as important to understanding dementia severity and potential treatment effects. These considerations are motivating additional inquiries by our group.

Acknowledgments

DGI is supported by the National Research Council - Canada Industrial Research Assistance Program (NRC-IRAP), the Atlantic Canada Opportunities Agency (ACOA), and Innovacorp.

Conflicts of Interest

This study was sponsored by Janssen Alzheimer Immunotherapy R&D, LLC, and Pfizer Inc. Kenneth Rockwood, Matthew Richard, and Arnold Mitnitski are employees of DGI Clinical Inc and were paid consultants to Janssen Alzheimer Immunotherapy R&D, LLC, in connection with the development of this manuscript. Kenneth Rockwood is founder of DGI Clinical Inc (the company was then called DementiaGuide Inc) and has a proprietary interest. Lisa Mucha is an employee of Pfizer Inc. Christopher Leibman is an employee of Janssen Alzheimer Immunotherapy R&D, LLC.

Multimedia Appendix 1

SymptomGuide screenshots.

[[PPTX File, 1MB - jmir_v15i8e145_app1.pptx](#)]]

Multimedia Appendix 2

The online Symptom Library.

[[MP4 File \(MP4 Video\), 4MB - jmir_v15i8e145_app2.mp4](#)]]

References

1. Fox S. 80% of Internet users look for health information online.: Pew Internet & American Life Project; 2011. URL: http://www.pewinternet.org/~media/Files/Reports/2011/PIP_Health_Topics.pdf [accessed 2013-07-03] [[WebCite Cache ID 6Hq5EzT17](#)]

2. Chow TW, Pio FJ, Rockwood K. An international needs assessment of caregivers for frontotemporal dementia. *Can J Neurol Sci* 2011 Sep;38(5):753-757. [Medline: [21856580](#)]
3. Rockwood K, Zeng A, Leibman C, Mucha L, Mitnitski A. Validation of an informant-reported web-based data collection to assess dementia symptoms. *J Med Internet Res* 2012;14(2):e42 [FREE Full text] [doi: [10.2196/jmir.1941](#)] [Medline: [22411293](#)]
4. Chiu TM, Eysenbach G. Theorizing the health service usage behavior of family caregivers: a qualitative study of an internet-based intervention. *Int J Med Inform* 2011 Nov;80(11):754-764. [doi: [10.1016/j.ijmedinf.2011.08.010](#)] [Medline: [21958550](#)]
5. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 Internet-based intervention programs. *Am J Alzheimers Dis Other Demen* 2011 Feb;26(1):36-43. [doi: [10.1177/1533317510387586](#)] [Medline: [21282276](#)]
6. Chiu TM, Eysenbach G. Stages of use: consideration, initiation, utilization, and outcomes of an internet-mediated intervention. *BMC Med Inform Decis Mak* 2010;10:73 [FREE Full text] [doi: [10.1186/1472-6947-10-73](#)] [Medline: [21092275](#)]
7. Chiu T, Marziali E, Colantonio A, Carswell A, Gruneir M, Tang M, et al. Internet-based caregiver support for Chinese Canadians taking care of a family member with alzheimer disease and related dementia. *Can J Aging* 2009 Dec;28(4):323-336. [doi: [10.1017/S0714980809990158](#)] [Medline: [19925698](#)]
8. Kernisan LP, Sudore RL, Knight SJ. Information-seeking at a caregiving website: a qualitative analysis. *J Med Internet Res* 2010;12(3):e31 [FREE Full text] [doi: [10.2196/jmir.1548](#)] [Medline: [20675292](#)]
9. Stern Y, Albert SM, Sano M, Richards M, Miller L, Folstein M. Assessing patient dependence in Alzheimer's disease. *J Gerontol* 1994;49:M216-M222.
10. Rockwood K. An individualized approach to tracking and treating Alzheimer's disease. *Clin Pharmacol Ther* 2010 Oct;88(4):446-449. [doi: [10.1038/clpt.2010.68](#)] [Medline: [20856240](#)]
11. Kecman V. *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*. Cambridge, MA: MIT Press; 2001.
12. Quintana M, Guàrdia J, Sánchez-Benavides G, Aguilar M, Molinuevo JL, Robles A, Neuronorma Study Team. Using artificial neural networks in clinical neuropsychology: high performance in mild cognitive impairment and Alzheimer's disease. *J Clin Exp Neuropsychol* 2012;34(2):195-208. [doi: [10.1080/13803395.2011.630651](#)] [Medline: [22165863](#)]
13. Tabaton M, Odetti P, Cammarata S, Borghi R, Monacelli F, Caltagirone C, et al. Artificial neural networks identify the predictive values of risk factors on the conversion of amnesic mild cognitive impairment. *J Alzheimers Dis* 2010;19(3):1035-1040. [doi: [10.3233/JAD-2010-1300](#)] [Medline: [20157257](#)]
14. Grossi E, Buscema MP, Snowdon D, Antuono P. Neuropathological findings processed by artificial neural networks (ANNs) can perfectly distinguish Alzheimer's patients from controls in the Nun Study. *BMC Neurol* 2007;7:15 [FREE Full text] [doi: [10.1186/1471-2377-7-15](#)] [Medline: [17584929](#)]
15. Licastro F, Porcellini E, Chiappelli M, Forti P, Buscema M, Ravaglia G, et al. Multivariable network associated with cognitive decline and dementia. *Neurobiol Aging* 2010 Feb;31(2):257-269. [doi: [10.1016/j.neurobiolaging.2008.03.019](#)] [Medline: [18485535](#)]
16. Song X, Mitnitski A, MacKnight C, Rockwood K. Assessment of individual risk of death using self-report data: an artificial neural network compared with a frailty index. *J Am Geriatr Soc* 2004 Jul;52(7):1180-1184. [doi: [10.1111/j.1532-5415.2004.52319.x](#)] [Medline: [15209659](#)]
17. Song X, Mitnitski A, Cox J, Rockwood K. Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *Stud Health Technol Inform* 2004;107(Pt 1):736-740. [Medline: [15360910](#)]
18. Franceschi M, Caffarra P, Savarè R, Cerutti R, Grossi E, Tol Research Group. Tower of London test: a comparison between conventional statistic approach and modelling based on artificial neural network in differentiating fronto-temporal dementia from Alzheimer's disease. *Behav Neurol* 2011;24(2):149-158. [doi: [10.3233/BEN-2011-0327](#)] [Medline: [21606576](#)]
19. Symptom Library. URL: <http://www.dementiaguide.com/symptomlibrary/> [accessed 2013-07-03] [WebCite Cache ID [6Hq5hXvpm](#)]
20. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982 Sep;139(9):1136-1139. [Medline: [7114305](#)]
21. Reisberg B, Ferris SH. Brief Cognitive Rating Scale (BCRS). *Psychopharmacol Bull* 1988;24(4):629-636. [Medline: [3249764](#)]
22. Song X, Mitnitski A, Rockwood K. Nontraditional risk factors combine to predict Alzheimer disease and dementia. *Neurology* 2011 Jul 19;77(3):227-234 [FREE Full text] [doi: [10.1212/WNL.0b013e318225c6bc](#)] [Medline: [21753161](#)]

Abbreviations

- ANOVA:** analysis of variance
- ANN:** artificial neural network
- CIND:** cognitive impairment no dementia
- MCI:** mild cognitive impairment
- MMSE:** Mini-Mental State Examination

Edited by G Eysenbach; submitted 03.12.12; peer-reviewed by S López-Pousa, G Smith, A Jansen; comments to author 14.02.13; revised version received 25.03.13; accepted 10.06.13; published 07.08.13.

Please cite as:

Rockwood K, Richard M, Leibman C, Mucha L, Mitnitski A

Staging Dementia From Symptom Profiles on a Care Partner Website

J Med Internet Res 2013;15(8):e145

URL: <http://www.jmir.org/2013/8/e145/>

doi: [10.2196/jmir.2461](https://doi.org/10.2196/jmir.2461)

PMID: [23924608](https://pubmed.ncbi.nlm.nih.gov/23924608/)

©Kenneth Rockwood, Matthew Richard, Chris Leibman, Lisa Mucha, Arnold Mitnitski. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 07.08.2013. 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

Internet-Based Early Intervention to Prevent Posttraumatic Stress Disorder in Injury Patients: Randomized Controlled Trial

Joanne Mouthaan¹, MSc; Marit Sijbrandij^{2,3}, PhD; Giel-Jan de Vries¹, MSc; Johannes B Reitsma^{4,5}, MD, PhD; Rens van de Schoot^{6,7}, PhD; J Carel Goslings⁸, MD, PhD; Jan SK Luitse⁸, MD; Fred C Bakker⁹, MD, PhD; Berthold PR Gersons^{1,10}, MD, PhD; Miranda Olf^{1,10}, PhD

¹Center for Anxiety Disorders, Research Group Psychotrauma, Department of Psychiatry, Academic Medical Center, Amsterdam, Netherlands

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

³EMGO Institute for Health and Care Research, Amsterdam, Netherlands

⁴Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, Amsterdam, Netherlands

⁵Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht, Netherlands

⁶Department of Methodology and Statistics, Utrecht University, Utrecht, Netherlands

⁷Optentia Research Program, Faculty of Humanities, North-West University, Potchefstroom, South Africa

⁸Trauma Unit, Department of Surgery, Academic Medical Center, Amsterdam, Netherlands

⁹Department of Traumatology, VU University Medical Center, Amsterdam, Netherlands

¹⁰Arq Psychotrauma Expert Group, Diemen, Netherlands

Corresponding Author:

Joanne Mouthaan, MSc

Center for Anxiety Disorders, Research Group Psychotrauma

Department of Psychiatry

Academic Medical Center

Meibergdreef 5

Amsterdam, 3356 BE

Netherlands

Phone: 31 208913552

Fax: 31 208913664

Email: j.mouthaan@amc.uva.nl

Abstract

Background: Posttraumatic stress disorder (PTSD) develops in 10-20% of injury patients. We developed a novel, self-guided Internet-based intervention (called Trauma TIPS) based on techniques from cognitive behavioral therapy (CBT) to prevent the onset of PTSD symptoms.

Objective: To determine whether Trauma TIPS is effective in preventing the onset of PTSD symptoms in injury patients.

Methods: Adult, level 1 trauma center patients were randomly assigned to receive the fully automated Trauma TIPS Internet intervention (n=151) or to receive no early intervention (n=149). Trauma TIPS consisted of psychoeducation, in vivo exposure, and stress management techniques. Both groups were free to use care as usual (nonprotocolized talks with hospital staff). PTSD symptom severity was assessed at 1, 3, 6, and 12 months post injury with a clinical interview (Clinician-Administered PTSD Scale) by blinded trained interviewers and self-report instrument (Impact of Event Scale—Revised). Secondary outcomes were acute anxiety and arousal (assessed online), self-reported depressive and anxiety symptoms (Hospital Anxiety and Depression Scale), and mental health care utilization. Intervention usage was documented.

Results: The mean number of intervention logins was 1.7, SD 2.5, median 1, interquartile range (IQR) 1-2. Thirty-four patients in the intervention group did not log in (22.5%), 63 (41.7%) logged in once, and 54 (35.8%) logged in multiple times (mean 3.6, SD 3.5, median 3, IQR 2-4). On clinician-assessed and self-reported PTSD symptoms, both the intervention and control group showed a significant decrease over time ($P<.001$) without significant differences in trend. PTSD at 12 months was diagnosed in 4.7% of controls and 4.4% of intervention group patients. There were no group differences on anxiety or depressive symptoms over time. Post hoc analyses using latent growth mixture modeling showed a significant decrease in PTSD symptoms in a subgroup of patients with severe initial symptoms (n=20) ($P<.001$).

Conclusions: Our results do not support the efficacy of the Trauma TIPS Internet-based early intervention in the prevention of PTSD symptoms for an unselected population of injury patients. Moreover, uptake was relatively low since one-fifth of individuals did not log in to the intervention. Future research should therefore focus on innovative strategies to increase intervention usage, for example, adding gameplay, embedding it in a blended care context, and targeting high-risk individuals who are more likely to benefit from the intervention.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 57754429; <http://www.controlled-trials.com/ISRCTN57754429> (Archived by WebCite at <http://webcitation.org/6FeJtJJyD>).

(*J Med Internet Res* 2013;15(8):e165) doi:[10.2196/jmir.2460](https://doi.org/10.2196/jmir.2460)

KEYWORDS

early intervention; prevention; Internet; posttraumatic stress disorder; cognitive behavior therapy

Introduction

Posttraumatic stress disorder (PTSD) develops after trauma exposure, such as violence, disasters, and injury [1,2]. PTSD's lifetime prevalence in adults is 7-8% [3,4], whereas the conditional prevalence rate after exposure to violence or injury ranges from 10-56% [1,3,5]. PTSD symptoms include intrusions of the traumatic event, avoidance of stimuli related to the event, emotional numbness, and hyperarousal [6]. Until now, efforts to prevent PTSD onset, for example, psychological debriefing, have been unsuccessful [7,8]. Early treatment of PTSD, or its precursor Acute Stress Disorder, with 4-5 sessions of trauma-focused cognitive behavioral therapy (CBT) was found to be effective in preventing chronic PTSD [9]. CBT consists of imaginal exposure to the traumatic incident, aimed at extinction of the original fear associations [10], and stress-management techniques and cognitive restructuring to correct irrational beliefs [11]. A recent randomized controlled trial found evidence for the effectiveness of 3 sessions of prolonged (imaginal) exposure, starting within 12 hours of the traumatic event, in counteracting later symptoms of PTSD and depression [12]. It is yet unclear whether CBT-techniques administered as a single session early intervention are effective in preventing PTSD.

We developed Trauma TIPS, a brief self-guided Internet intervention based on established CBT techniques. Trauma TIPS aims to decrease acute levels of distress, anxiety, and arousal, known to predict PTSD [13], and to prevent the onset of PTSD symptoms by providing information on successful coping, instructions for self-exposure to fearful situations, and stress management techniques. The exponential growth of global Internet use contributes to the feasibility of e-mental health interventions, which are considered a cost-effective alternative to traditional interventions [14]. Although both self-guided and therapist-assisted Internet-based CBT programs have been successful in the *treatment* of PTSD [15], there is a great lack of study into whether these programs may *prevent* PTSD. Preliminary evidence from one previous study on the efficacy of a self-guided Internet-based psychoeducational program for injured children and their parents showed greater anxiety reductions in children who had completed the program compared to those who had not [16].

Our study examined whether Trauma TIPS prevents the onset of PTSD symptoms in injury patients compared to care as usual. In addition, we evaluated whether Trauma TIPS prevented

symptoms of depression and anxiety and led to a decrease in mental health care utilization during the first year after injury.

Methods

Trial Design

This study was an assessor-blinded randomized controlled trial (RCT; ISRCTN57754429) comparing a brief Internet-based early psychological intervention with a care-as-usual control group in two trauma centers (see [Multimedia Appendix 1](#) for the CONSORT E-HEALTH Checklist of the trial).

Participants

Injury patients transported by ambulance or helicopter to the level 1 trauma centers of the Academic Medical Center (AMC) and VU University Medical Center (VUmc) in Amsterdam, the Netherlands, were eligible for inclusion. These patients were suspected to suffer from possible severe injuries that required specialized acute medical care. Inclusion criteria were age 18 years or older, proficiency in Dutch, and having experienced a potential traumatic event (cf. Criterion A1 DSM-IV PTSD diagnosis) [6]. According to this criterion, the person has experienced, witnessed, or been confronted with an event or events that involve actual or threatened death or serious injury, or a threat to the physical integrity of oneself or others. Exclusion criteria were the injury resulting from deliberate self-harm; organic brain condition, psychotic disorder, bipolar disorder, or depression with psychotic features (cf. DSM-IV) [6]; moderate to severe traumatic brain injury (TBI) (according to a Glasgow Coma Score [17] less than 13); and permanent residency outside the Netherlands.

Interventions

Trauma TIPS [18] (for screenshots see [Multimedia Appendices 2 and 3](#)) was created and is owned by the authors from the Research Group Psychotrauma [19]. It is based on CBT techniques of psychoeducation, stress management/relaxation techniques, and in vivo exposure. It consists of 6 steps, including introduction to the program and basic operating instructions; assessments of acute anxiety and arousal using Visual Analogue Scales (VAS) at pre- and postintervention; video features of the trauma center's surgical head explaining the procedures at the center and the purpose of the program, and of 3 patient models sharing their experiences after their injury; a short textual summary of 5 coping tips for common physical and psychological reactions after trauma; audio clips with

instructions for stress management techniques; contact information for program assistance or professional help for enduring symptoms; and a Web forum for peer support. The introduction page shows the logos of the academic hospitals involved in the study, as well as the logos of the funders of the study. The full design and content of the intervention are described elsewhere [19,20]. Total duration of the program was approximately 30 minutes. Care as usual, available to patients from both groups, consisted of incidental, nonstructured talks with trauma center staff or with a patient's general practitioner (GP), either directly following injury or during the course of the trial.

Study Procedures

The local institutional review boards provided medical ethical approval. Patients were contacted in hospital or via telephone within 72 hours post injury to assess eligibility and to schedule a baseline assessment. Informed consent was obtained face-to-face directly prior to the baseline assessment at approximately 1 week post injury. Patients were randomly allocated to (1) the Trauma TIPS intervention or (2) a control group with no intervention, but access to care as usual. Randomization was performed by a research member independent of data collection in a 1:1 ratio by a computerized program, TENALEA Clinical Trial Data Management System (NKI/AVL Biometrics department, Amsterdam), using random block sizes (with maximum block size 6), stratified by study center. Intervention group patients received personal log-in codes for the intervention's website, along with instructions to perform the intervention at will, but at least once within the first month. Electronic and telephone reminders were sent to encourage (early) log-in, but patients were free to access the intervention as they pleased, to underscore the intervention's voluntary nature and self-guiding principles. Research assistants visited patients with a laptop in case of hospitalization or a lack of Internet or computer access. Follow-up assessments were scheduled at 1, 3, 6, and 12 months post injury. The assessments took place at the AMC's Center for Anxiety Disorders, at bedside in the hospital, or at the private home of the patient. Patients were asked not to share information about the randomization to the assessors, to ensure that they were blind to the allocated interventions. No reimbursement was given.

Outcomes

Trained assessors at the master's and doctoral levels performed the data collection. The main outcome measure was PTSD symptom severity on the Clinician-Administered PTSD Scale (CAPS) [21]. The structured interview assesses the frequency and intensity (ranging from 0-4) of the 17 DSM IV symptoms of PTSD (total scores range from 0-136). Scores are added to represent PTSD symptom severity or a diagnosis. The internal consistency of the Dutch translation of the CAPS is good to excellent [22]. Presence of a PTSD diagnosis was computed using the established rule of Weathers et al [23].

The Mini International Neuropsychiatric Interview (MINI-Plus, version 5.0) [24], a semistructured clinical interview, was used to obtain DSM IV diagnoses of major depressive disorders (MDD) and anxiety disorders other than PTSD. Each module

starts with screening questions, which, if positive, lead to a further examination of the disorder's criteria.

We assessed self-reported PTSD severity with the Impact of Event Scale-Revised (IES-R) [25]. The 22 items are scored on a 5-point scale, from 0 (not at all) to 4 (extremely). Total scores range from 0-88 with higher scores representing more severe symptoms. The IES-R shows high internal consistency [25,26].

Self-reported severity of depressive and anxiety symptoms was assessed using the Hospital Anxiety and Depression Scale (HADS) [27]. The item scores in the two subscales of depression (7 items) and anxiety (7 items) range from 0-3 (total scores per subscale ranging from 0-21). Higher scores indicate greater symptomatology. The test-retest reliability of the 2 scales is high [28].

The Trimbos/iMTA questionnaire for Costs associated with Psychiatric illness (TiC-P) [29] was used to evaluate direct and indirect health costs. Direct costs include contacts with mental health professionals (eg, GP, psychologist, social worker), medication use, and admissions for mental health problems. Indirect costs were calculated as production losses due to psychological problems by the Short Form Health and Labour Questionnaire (SF-HLQ) [30].

At the beginning and after completion of Trauma TIPS, patients indicated acute anxiety and arousal levels from 0 (no anxiety or arousal) to 100 (worst anxiety or arousal) on two online VASs [19,20].

Website activity was recorded to evaluate usage characteristics, such as number of log-ins and total amount of login time.

Sample Size

To demonstrate a difference of at least 5.5 points on the CAPS between the groups at 12 months, equivalent to a small to medium effect size of Cohen's $d=.35$, 134 patients or more per condition were required (Cronbach alpha=.05, power=80%, SD 16) [31]. Anticipating possible attrition of study participants, we included 150 patients per condition.

Analyses

Differences in baseline characteristics between the study groups, patients lost to follow-up vs patients not lost to follow-up and patient groups with varying intervention usage were tested using independent sample t tests and chi-square tests (Bonferroni adjusted $P=.005$). Missing data were imputed using general purpose multivariate imputation procedure (ICE: sequential regression imputation method), creating 50 different datasets. All analyses were performed using these 50 datasets and then pooled by combining the individual results. Due to their positive skewness, CAPS and IES-R values were square root transformed. Stata version 11.2 was used for all repeated measures analyses of PTSD symptoms (CAPS, IES-R) and depressive and anxiety symptoms (HADS-A, HADS-D). The effects of time of measurement, group, and the group-by-time interaction were analyzed with linear mixed models. For all regression models, a robust variance estimator was used. Estimated values (adjusted) and 95% confidence intervals (CIs) are presented throughout the paper unless otherwise specified. Finally, as a post hoc analysis, we applied latent growth mixture

modelling (LGMM) [32,33] to explore possible latent subgroups within the two groups by use of the software Mplus (Version 6.11) [34] using a Bayesian estimator [35,36]. Across all analyses, two-tailed tests are reported with Cronbach alpha=.05.

Results

Baseline Characteristics

Recruitment and follow-up took place from September 2007 to June 2010. Figure 1 shows the flow of patients through the trial. Participants were significantly older (mean age 43.8, SD 15.9) than patients who refused participation (mean age 40.1, SD 16.3, $P=.01$). Table 1 shows the baseline characteristics of participants. There were no differences in baseline characteristics or attrition rate between the study groups. Patients lost to the 12-month follow-up were more often unmarried than patients who were not lost to follow-up ($P=.001$).

Intervention Usage

Most intervention group patients logged in to the intervention's website once ($n=63$, 41.7%). Fifty-four patients (35.8%) logged in multiple times (mean 3.6, SD 3.5, median 3, IQR 2-4). Thirty-four patients (22.5%) did not log in (ie, nonusers) and provided the following reasons: not interested anymore (2), occupied with rehabilitation (1), too busy (1), on holiday (1), too much on my mind (1), tired (1), difficulty concentrating (1), postconcussion symptoms (1), broken back (1), husband deceased (1), or no explanation (22). The average number of log-ins for the entire group was 1.7 (SD 2.5). The average login time was 20.8 minutes (SD 26.3). There were no differences in attrition or outcome measures between nonusers ($n=34$) and users of the intervention ($n=117$), or between patients with a single log-in ($n=63$) versus multiple log-ins ($n=54$). The only differences were that more nonusers than users had a non-Dutch cultural background ($P=.003$) and that patients with multiple log-ins were significantly older (mean age 48.0, SD 14.6) than those with a single log-in (mean age 39.6, SD 14.1, $P=.001$).

From pre- to postintervention, the majority of intervention group patients reported no change in acute anxiety (55.9%, $n=38$) and arousal (63.2%, $n=43$) on the VASs. Seven patients reported an increase (10.3%), and 23 (33.8%) and 18 (26.5%) patients reported a reduction in anxiety and arousal respectively.

Main Outcomes

Table 2 shows the results of the intention-to-treat analyses for PTSD, anxiety, and depressive symptoms. Mixed-model analysis of PTSD symptom severity of the CAPS showed a significant effect of time ($P<.001$), but no significant group differences over time (12-month follow-up, Internet intervention group: estimated means 13.0, 95% CI 11.2 - 14.8; control group:

estimated means 13.0, 95% CI 11.4 - 14.6, $P=.63$). On the mixed-model analysis of self-reported PTSD symptoms (IES-R), we found a similar significant time effect ($P<.001$) and no group differences over time (12-month follow-up, Internet intervention group: estimated means 7.6, 95% CI 6.4 - 8.7; control group: estimated means 7.8, 95% CI 6.4 - 9.2, $P=.76$). Figure 2 presents the estimated CAPS and IES-R means over time. For depressive and anxiety symptoms, we found no effects of time or group over time in mixed-model analyses (12 month HADS-D, Internet intervention group: estimated means 3.3, 95% CI 2.4 - 4.2; control group: estimated means 3.0, 95% CI 2.2 - 3.7, $P=.72$; 12 month HADS-A, Internet intervention group: estimated means 4.1, 95% CI 3.5 - 4.8; control group: estimated means 3.7, 95% CI 3.0 - 4.3, $P=.53$).

PTSD was diagnosed in 9.2% of patients at 1 month ($n=21$), 7.6% at 3 months ($n=14$), 7.5% at 6 months ($n=11$), and 4.5% at 12 months ($n=6$). MDD was diagnosed in 7.6% of patients at 1 month ($n=17$), 2.7% at 3 months ($n=5$), 7.6% at 6 months ($n=11$), and 6.8% at 12 months ($n=9$). Ten patients (4.4%) were diagnosed with an anxiety disorder at 1 month, 11 patients (6.0%) at 3 months, 14 patients (9.7%) at 6 months, and 10 patients (7.6%) at 12 months. chi-square analyses showed no group differences in prevalence of any of the psychiatric diagnoses.

Mental health care utilization at 12 months was similar for both groups, such as visits to a GP ($P=.35$), company doctor ($P=.95$), mental health specialists ($P=.52$), hospital admissions ($P=.70$), or medication use ($P=.57$). The groups also did not differ with respect to employment status ($P=.70$), working hours ($P=.89$), and work absence ($P=.81$). Due to the absence of significant group differences, the direct and indirect costs for mental health use were not calculated.

Completer Analyses

In completers-only analyses ($n=117$ intervention group and $n=149$ control group patients), excluding nonusers ($n=34$), results were similar to the intention-to-treat results for all outcome measures.

Latent Subgroups

Post hoc LGMM analyses of self-reported PTSD symptoms (IES-R) revealed two latent subgroups per study group based on PTSD symptom severity at baseline, resulting in a low symptomatic control subgroup ($n=94$) and intervention subgroup ($n=105$), and a high symptomatic control subgroup ($n=15$) and intervention subgroup ($n=20$). The main difference between the groups was the slope of the high symptomatic subgroups, which showed a significant decrease in the intervention subgroup ($P<.001$), but not in the control subgroup ($P=.32$). Table 3 shows the outcomes of the LGMM analyses.

Table 1. Participant characteristics at baseline.

Characteristic	Internet intervention n=151	Control with usual care n=149	<i>P</i> value ^a
Age in years, mean (SD)	44.18 (15.76)	43.49 (16.00)	.54
Sex (male), n (%)	89 (58.9)	91 (61.1)	.73
Post-high school education, n (%)	37 (24.7)	43 (29.1)	.71
Unemployed, n (%)	41 (27.5)	29 (19.5)	.13
Married/cohabitating, n (%)	82 (54.3)	81 (54.4)	.54
Dutch cultural background, n (%)	127 (84.1)	122 (83.0)	.88
Prior traumatic events, mean (SD)	2.99 (2.42)	2.93 (2.20)	.80
Hospital admission, n (%)	100 (66.7)	105 (70.9)	.46
Days hospitalized, mean (SD)	5.30 (8.02)	4.57 (7.36)	.20
ICU admission, n (%)	13 (8.7)	13 (8.8)	.97
Injury Severity Score, mean (SD)	10.45 (8.59)	10.21 (9.87)	.33
Glasgow Coma Scale, mean (SD)	14.48 (1.91)	14.72 (1.42)	.08
Traumatic event, n (%)			.11
Traffic accident	99 (65.6)	105 (70.5)	
Work-related accident	12 (7.9)	16 (10.7)	
Fall	28 (18.5)	13 (8.7)	
Interpersonal violence/physical abuse	2 (2.3)	5 (3.4)	
Other	10 (6.6)	10 (6.7)	
Psychological assessment tools, mean (SD)			
Impact of Event Scale—Revised	17.60 (16.82)	21.22 (19.09)	.15
Hospital Anxiety and Depression Scale—Depression	3.69 (3.50)	4.13 (4.26)	.09
Hospital Anxiety and Depression Scale—Anxiety	4.36 (3.90)	4.87 (4.33)	.21

^aIndependent *t* test for difference between groups for continuous measures and chi-square test for differences between groups in categorical characteristics.

Figure 1. Flow of participants through the trial.

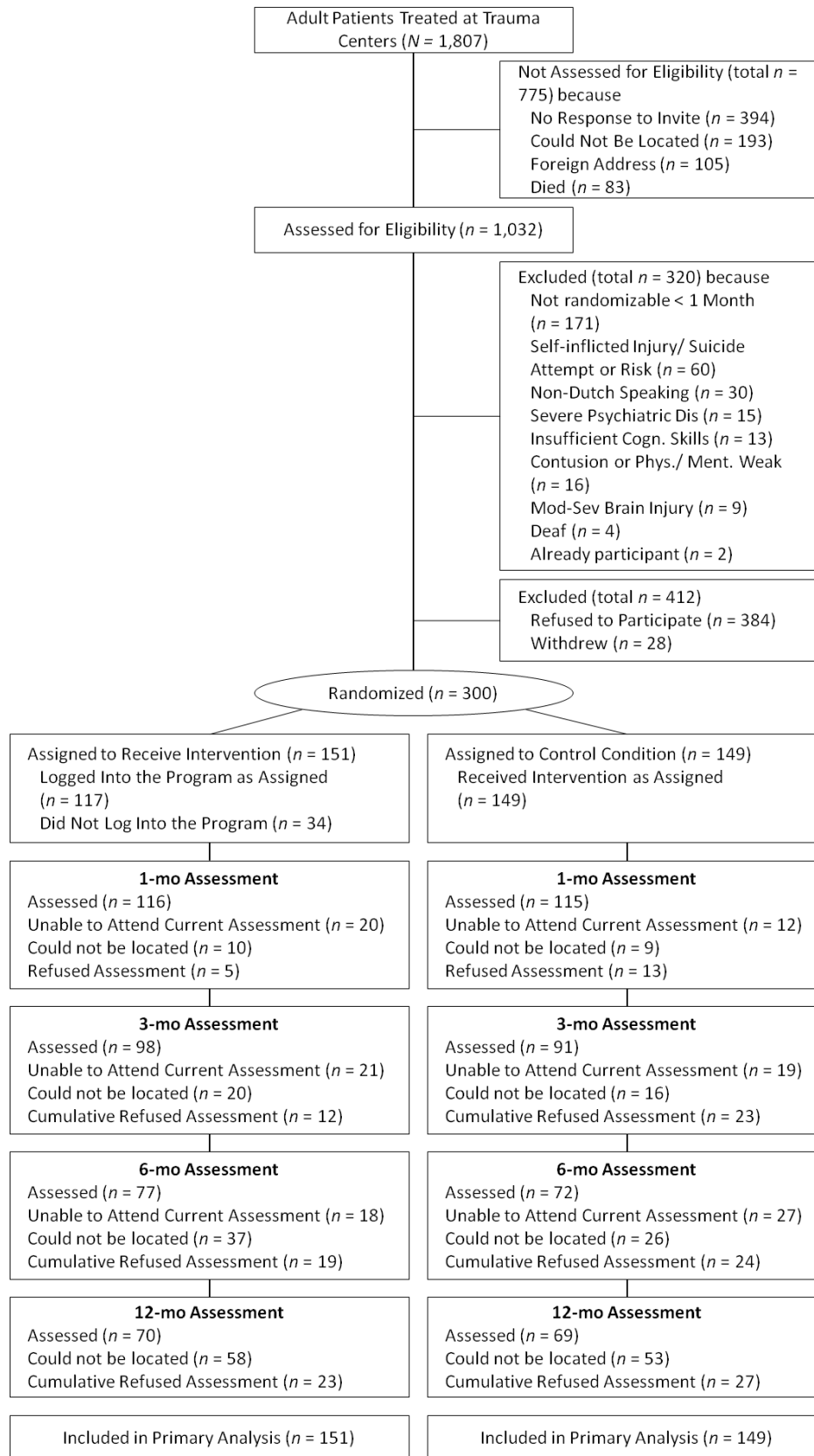


Table 2. Outcomes of intention-to-treat linear mixed models for PTSD, depressive, and anxiety symptoms.^a

Outcome	Internet intervention	Control with usual care n=149	Time		Group		Group x Time	
			<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>	<i>F</i>	<i>P</i>
Clinician-assessed PTSD symptoms (CAPS)			6.3	<.001	1.7	.19	0.6	.63
1 month follow-up	17.7 (16.7 to 18.7)	20.2 (19.1 to 21.3)						
3 month follow-up	14.3 (13.2 to 15.5)	16.8 (15.1 to 18.6)						
6 month follow-up	14.5 (13.2 to 15.8)	15.7 (14.3 to 17.1)						
12 month follow-up	13.0 (11.2 to 14.8)	13.0 (11.4 to 14.6)						
Patient-reported PTSD symptoms (IES-R)			15.7	<.001	1.9	.17	0.5	.76
1 month follow-up	10.6 (9.6 to 11.7)	12.4 (11.1 to 13.7)						
3 month follow-up	9.7 (8.0 to 11.4)	11.8 (10.1 to 13.5)						
6 month follow-up	8.2 (6.9 to 9.6)	9.8 (8.1 to 11.5)						
12 month follow-up	7.6 (6.4 to 8.7)	7.8 (6.4 to 9.2)						
Anxiety symptoms (HADS-A)			2.2	.07	0.3	.57	0.8	.53
1 month follow-up	4.6 (3.9 to 5.2)	4.8 (4.1 to 5.5)						
3 month follow-up	4.0 (3.5 to 4.5)	4.3 (3.8 to 4.9)						
6 month follow-up	3.9 (3.2 to 4.6)	4.6 (3.7 to 5.4)						
12 month follow-up	4.1 (3.4 to 4.8)	3.7 (3.0 to 4.3)						
Depressive symptoms (HADS-D)			2.3	.054	0.3	.62	0.5	.72
1 month follow-up	3.6 (3.2 to 4.0)	4.1 (3.5 to 4.6)						
3 month follow-up	3.5 (3.0 to 4.0)	3.9 (3.4 to 4.5)						
6 month follow-up	4.1 (3.5 to 4.8)	4.5 (3.6 to 5.4)						
12 month follow-up	3.3 (2.4 to 4.2)	3.0 (2.2 to 3.7)						

^aData are expressed as mean (95% CI).

Figure 2. Trends in observed PTSD symptom severity (CAPS and IES-R) per intervention group.

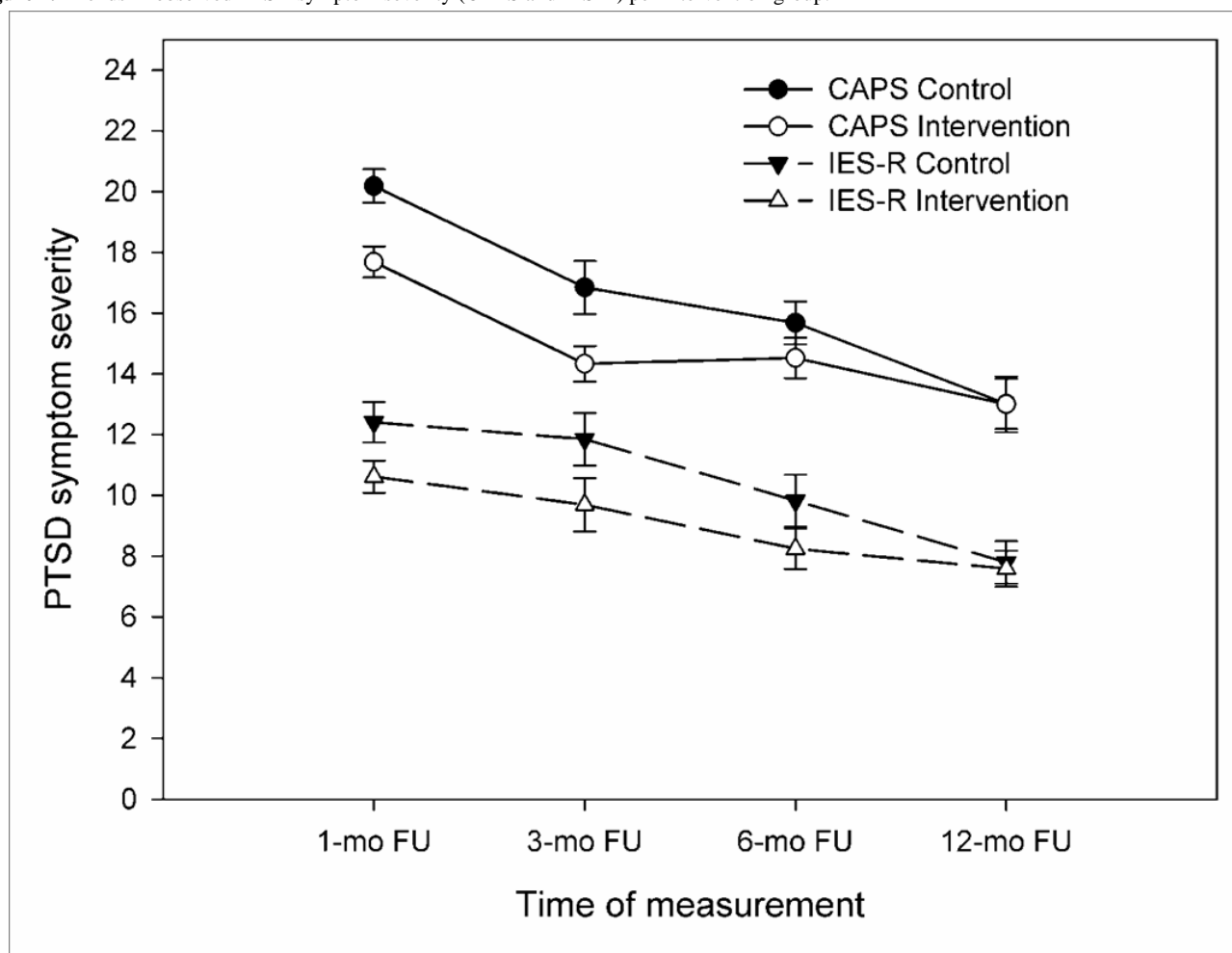


Table 3. Outcomes of latent growth mixture modeling analyses for self-reported PTSD severity (IES-R).

Latent subgroups	Internet intervention			Control with usual care		
	n	mean (95% CI)	P	n	mean (95% CI)	P
Low symptomatic subgroup	105			94		
Intercept		9.0 (6.9 to 11.1)	<.001		14.9. (11.4 to 18.5)	<.001
Slope		-1.0 (-1.4 to -0.4)	<.001		-1.4 (-1.9 to -0.8)	<.001
High symptomatic subgroup	20			15		
Intercept		41.2 (35.0 to 50.3)	<.001		42.9 (30.1 to 55.6)	<.001
Slope		-3.6 (-5.2 to -2.1)	<.001		0.6 (-2.7 to 3.9)	.32

Discussion

Principal Findings

In this paper, we presented the results of a randomized clinical trial comparing a self-guided Internet-based prevention program vs usual care in the prevention of PTSD symptoms in injury patients. PTSD symptoms decreased over time without a significant difference between the Internet intervention group and the control group. Moreover, there were no differences between groups with respect to the number of PTSD and MDD diagnoses and with respect to the severity of depression and anxiety at 12 months. An important finding is that participants

were reluctant to use the intervention. In fact, one in five patients in the intervention group lacked any exposure to the intervention. Based on these results, there are currently no indications that offering a voluntary, information-based prevention program via the Internet to unselected injury victims is useful in preventing PTSD symptoms.

The low adherence rates were comparable to those found in similar self-help Internet-based interventions [37]. In part, this nonusage was a consequence of a deliberate design choice to allow patients freedom in performing the intervention, having learned from adverse effects of debriefing interventions found previously to be noneffective or even harmful [7,13]. However, in order to induce changes in behavior and affect, true exposure

to an intervention is necessary, which entails accessing the intervention website, staying on the intervention website to actually use it, and revisiting the intervention website, in case of a repetitive design [38]. As possible reasons for dropping out of or not adhering to online treatment programs, previous studies reported time constraints, lack of motivation, technical or computer-access problems, depressive episode or physical illness, the lack of face-to-face contact, a preference for taking medication, perceived lack of treatment effectiveness, improvement in condition, and burden of the program [37]. Strategies to increase uptake of Trauma TIPS may be a more structured peer-support forum, more interactive elements to the intervention, such as quizzes or knowledge questions, automated feedback on the acute anxiety and arousal assessments, or monetary incentive [38-40]. Moreover, a more strict approach to intervention adherence for inclusion in our study (eg, a minimum number of log-ins or log-in time required for participation) may have resulted in greater benefits. However, note that we found no differences in outcomes between users and nonusers or between participants with single versus multiple use. Finally, it is possible that the idea of a computerized program did not match the acute needs of the injury victims, resulting in some of them not using it. Previous studies investigating needs of victims after the September 11, 2001, terrorist attacks and the 2005 London bombings showed that only very few people (< 1%) reported a need for professional mental health support in the acute posttrauma phase, and most (71-87%) turned to loved ones or others for support [41,42].

Another explanation for not finding a significant effect of the intervention may be the low overall PTSD symptom level. Only 9.2% of patients developed PTSD at 1 month, which decreased to 4.5% at 12 months. Beforehand, we expected that 19% of participants would have developed PTSD at 3 months [31]. This unexpectedly low PTSD incidence left little room for symptom improvement for the whole group. Additionally, the relatively low symptom levels may have caused participants to experience little personal incentive to access and use the intervention. Support for this comes from our post hoc subgroup analyses that suggested that the Trauma TIPS intervention was effective in reducing PTSD symptoms in individuals with high initial symptom levels. Because this subgroup was small (n=20), these results must be interpreted with caution.

Internet interventions may not be suitable for all individuals. Common points of criticism are that the mainly information-driven formats pose a disadvantage to people with lesser reading or language skills, do not meet the needs of the elderly or persons with limited computer skills or experience, and that it is difficult to appeal to a culturally diverse audience

in a single format, as possibly illustrated in our sample of more nonusers having a non-Dutch cultural background [37]. On the other hand, the rapid developments in Internet applications, especially via mobile technology, provide more possibilities to reach populations who were earlier underserved in eHealth care [43,44].

Limitations

One limitation of our study was missing data due to patient dropout or failure to complete self-report instruments. We do not know to what extent attrition may have biased our results, although besides marital status, we found no differences between participants and dropouts. In addition, our sample may not have been fully representative of the entire level 1 trauma center population, since we excluded patients with moderate-severe TBI, who did not master the Dutch language, or who were unable to meet our time requirements for logging in.

Conclusions

As a clinical implication of our study, future comparable Internet-based early interventions should be aimed at individuals with high initial symptoms. These individuals may be accurately identified within the first weeks following trauma with early screening tools for PTSD [45-47]. Stepped care programs for acutely traumatized individuals have recently shown to be feasible [48]. The results of our study show that an e-mental health approach could well be a first step in the acute aftercare of highly distressed trauma victims, since Trauma TIPS was indeed effective in a latent subgroup of participants experiencing high levels of PTSD symptoms at baseline. For those victims whose symptoms remain, our self-guided early intervention could be followed by more specialized or traditional curative face-to-face treatment as part of a blended care strategy [49].

Future studies may determine the effectiveness of applying interventions such as Trauma TIPS to individuals with high levels of distress. They may also evaluate whether incorporation of strategies to increase adherence, for instance a motivating interviewing module or increasing the fun by adding serious gaming components to Trauma TIPS, may increase its effectiveness.

In conclusion, our study found no evidence for preventing the development of PTSD symptoms by offering a voluntary, information-based prevention program via the Internet to unselected injury trauma victims. Future research may focus on innovative strategies to increase intervention usage and targeting high-risk individuals who are more likely to benefit from the intervention.

Acknowledgments

We acknowledge the contributions of Susanne van Buschbach, MSc, Nina A. van der Togt, MSc, and Hendriek B. Bolding, MSc, and the entire Trauma TIPS interview team for research assistance and data collection, and the participants in the study for their efforts in providing the study data. Funding for this study was provided by the Netherlands Organization for Health Research and Development (ZonMw), Grant no. 62300038, and by the Achmea Foundation for Victims and Society (Stichting Achmea Slachtoffer en Samenleving). The role of the funders in the study was limited to providing financial support to conduct the trial

and to oversee the general progress of the trial. Publication costs were covered by the Open Access Fund of the Netherlands Organization for Scientific Research (NWO).

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist V1.6.2 [42].

[PDF File (Adobe PDF File), 991KB - [jmir_v15i8e165_app1.pdf](#)]

Multimedia Appendix 2

Screenshot of the Trauma TIPS internet intervention: video feature of the trauma center's surgical head.

[JPG File, 62KB - [jmir_v15i8e165_app2.jpg](#)]

Multimedia Appendix 3

Screenshot of the Trauma TIPS internet intervention: video features of patient models.

[JPG File, 67KB - [jmir_v15i8e165_app3.jpg](#)]

References

1. Rees S, Silove D, Chey T, Ivancic L, Steel Z, Creamer M, et al. Lifetime prevalence of gender-based violence in women and the relationship with mental disorders and psychosocial function. *JAMA* 2011 Aug 3;306(5):513-521. [doi: [10.1001/jama.2011.1098](#)] [Medline: [21813429](#)]
2. O'Donnell ML, Creamer M, Pattison P, Atkin C. Psychiatric morbidity following injury. *Am J Psychiatry* 2004 Mar;161(3):507-514. [Medline: [14992977](#)]
3. de Vries GJ, Olf M. The lifetime prevalence of traumatic events and posttraumatic stress disorder in the Netherlands. *J Trauma Stress* 2009 Aug;22(4):259-267. [doi: [10.1002/jts.20429](#)] [Medline: [19645050](#)]
4. 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](#)] [Medline: [15939839](#)]
5. Bryant RA, O'Donnell ML, Creamer M, McFarlane AC, Clark CR, Silove D. The psychiatric sequelae of traumatic injury. *Am J Psychiatry* 2010 Mar;167(3):312-320. [doi: [10.1176/appi.ajp.2009.09050617](#)] [Medline: [20048022](#)]
6. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV. Washington, DC: American Psychiatric Association; 2000.
7. Rose S, Bisson J, Wessely S. A systematic review of single-session psychological interventions ('debriefing') following trauma. *Psychother Psychosom* 2003;72(4):176-184. [Medline: [12792122](#)]
8. De Silva M, Maclachlan M, Devane D, Desmond D, Gallagher P, Schnyder U, et al. Psychosocial interventions for the prevention of disability following traumatic physical injury. *Cochrane Database Syst Rev* 2009(4):CD006422 [[FREE Full text](#)] [doi: [10.1002/14651858.CD006422.pub3](#)] [Medline: [19821365](#)]
9. Roberts NP, Kitchiner NJ, Kenardy J, Bisson JI. Systematic review and meta-analysis of multiple-session early interventions following traumatic events. *Am J Psychiatry* 2009 Mar;166(3):293-301. [doi: [10.1176/appi.ajp.2008.08040590](#)] [Medline: [19188285](#)]
10. Foa EB, Kozak MJ. Emotional processing of fear: exposure to corrective information. *Psychol Bull* 1986 Jan;99(1):20-35. [Medline: [2871574](#)]
11. Foa EB, Riggs DS, Gershuny BS. Arousal, numbing, and intrusion: symptom structure of PTSD following assault. *Am J Psychiatry* 1995 Jan;152(1):116-120. [Medline: [7802101](#)]
12. Rothbaum BO, Kearns MC, Price M, Malcoun E, Davis M, Ressler KJ, et al. Early intervention may prevent the development of posttraumatic stress disorder: a randomized pilot civilian study with modified prolonged exposure. *Biol Psychiatry* 2012 Dec 1;72(11):957-963. [doi: [10.1016/j.biopsych.2012.06.002](#)] [Medline: [22766415](#)]
13. Sijbrandij M, Olf M, Reitsma JB, Carlier IV, Gersons BP. Emotional or educational debriefing after psychological trauma. Randomised controlled trial. *Br J Psychiatry* 2006 Aug;189:150-155 [[FREE Full text](#)] [doi: [10.1192/bjp.bp.105.021121](#)] [Medline: [16880485](#)]
14. Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. *Health Technol Assess* 2006 Sep;10(33):iii, xi-ixiv, 1 [[FREE Full text](#)] [Medline: [16959169](#)]

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. Cox CM, Kenardy JA, Hendrikz JK. A randomized controlled trial of a web-based early intervention for children and their parents following unintentional injury. *J Pediatr Psychol* 2010 Jul;35(6):581-592 [FREE Full text] [doi: [10.1093/jpepsy/jsp095](https://doi.org/10.1093/jpepsy/jsp095)] [Medline: [19906829](https://pubmed.ncbi.nlm.nih.gov/19906829/)]
17. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974 Jul 13;2(7872):81-84. [Medline: [4136544](https://pubmed.ncbi.nlm.nih.gov/4136544/)]
18. Trauma TIPS. URL: <http://www.traumatips.nl/> [accessed 2013-04-05] [WebCite Cache ID 6FeK0io9X]
19. Mouthaan J, Sijbrandij M, Reitsma JB, Gersons BP, Olf M. Internet-based prevention of posttraumatic stress symptoms in injured trauma patients: design of a randomized controlled trial. *Eur J Psychotraumatol* 2011;2:8294 [FREE Full text] [doi: [10.3402/ejpt.v2i0.8294](https://doi.org/10.3402/ejpt.v2i0.8294)] [Medline: [22893814](https://pubmed.ncbi.nlm.nih.gov/22893814/)]
20. Mouthaan J, Sijbrandij M, Reitsma JB, Luitse JSK, Goslings JC, Olf M. Trauma TIPS: an internet-based intervention to prevent posttraumatic stress disorder in injured trauma patients. *Journal of Cybertherapy and Rehabilitation* 2011;4(3):331-340.
21. Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. The development of a Clinician-Administered PTSD Scale. *J Trauma Stress* 1995 Jan;8(1):75-90. [Medline: [7712061](https://pubmed.ncbi.nlm.nih.gov/7712061/)]
22. Hovens JE, van der Ploeg HM, Klaarenbeek MT, Bramsen I, Schreuder JN, Rivero VV. The assessment of posttraumatic stress disorder: with the Clinician Administered PTSD Scale: Dutch results. *J Clin Psychol* 1994 May;50(3):325-340. [Medline: [8071438](https://pubmed.ncbi.nlm.nih.gov/8071438/)]
23. Weathers FW, Ruscio AM, Keane TM. Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment* 1999;11(2):124-133. [doi: [10.1037/1040-3590.11.2.124](https://doi.org/10.1037/1040-3590.11.2.124)]
24. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33;quiz 34. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
25. Weiss DS, Marmar CR. The Impact of Event Scale-Revised. In: Wilson JP, Keane TM, editors. *Assessing psychological trauma and PTSD*. New York: Guilford Press; 1997:399-411.
26. Creamer M, Bell R, Failla S. Psychometric properties of the Impact of Event Scale - Revised. *Behav Res Ther* 2003 Dec;41(12):1489-1496. [Medline: [14705607](https://pubmed.ncbi.nlm.nih.gov/14705607/)]
27. 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/)]
28. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997 Mar;27(2):363-370. [Medline: [9089829](https://pubmed.ncbi.nlm.nih.gov/9089829/)]
29. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemens B. Manual Trimbos/iMTA questionnaire for Costs associated with Psychiatric illness (TiC-P). Handleiding Trimbos/iMTA questionnaire for Costs associated with Psychiatric illness (TiC-P). Rotterdam, The Netherlands: Institute for Medical Technology Assessment, Erasmus University Rotterdam; 2002.
30. van Roijen L, Essink-Bot ML, Koopmanschap MA, Bonsel G, Rutten FF. Labor and health status in economic evaluation of health care. The Health and Labor Questionnaire. *Int J Technol Assess Health Care* 1996;12(3):405-415. [Medline: [8840661](https://pubmed.ncbi.nlm.nih.gov/8840661/)]
31. Conlon L, Fahy TJ, Conroy R. PTSD in ambulant RTA victims: a randomized controlled trial of debriefing. *J Psychosom Res* 1999 Jan;46(1):37-44. [Medline: [10088980](https://pubmed.ncbi.nlm.nih.gov/10088980/)]
32. Preacher KJ, Wichman AL, MacCallum RC, Briggs NE. *Latent Growth Curve Modeling*. London: Sage; 2008.
33. Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* 2000 Jun;24(6):882-891. [Medline: [10888079](https://pubmed.ncbi.nlm.nih.gov/10888079/)]
34. Muthén LK, Muthén BO. *Mplus User's Guide*, Sixth edition. Los Angeles: Muthén & Muthén; 2010. URL: <http://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v6.pdf> [accessed 2013-08-09] [WebCite Cache ID 6kTgTDA0]
35. Lynch SM. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer; 2007.
36. Elliott MR, Gallo JJ, Ten Have TR, Bogner HR, Katz IR. Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction. *Biostatistics* 2005 Jan;6(1):119-143 [FREE Full text] [doi: [10.1093/biostatistics/kxh022](https://doi.org/10.1093/biostatistics/kxh022)] [Medline: [15618532](https://pubmed.ncbi.nlm.nih.gov/15618532/)]
37. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009 Apr;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/)]
38. Crutzen R, Cyr D, de Vries NK. Bringing loyalty to e-Health: theory validation using three internet-delivered interventions. *J Med Internet Res* 2011 Sep;13(3):e73 [FREE Full text] [doi: [10.2196/jmir.1837](https://doi.org/10.2196/jmir.1837)] [Medline: [21946128](https://pubmed.ncbi.nlm.nih.gov/21946128/)]
39. 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/)]

40. 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 Aug;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/)]
41. Schuster MA, Stein BD, Jaycox L, Collins RL, Marshall GN, Elliott MN, et al. A national survey of stress reactions after the September 11, 2001, terrorist attacks. *N Engl J Med* 2001 Nov 15;345(20):1507-1512. [doi: [10.1056/NEJM200111153452024](https://doi.org/10.1056/NEJM200111153452024)] [Medline: [11794216](https://pubmed.ncbi.nlm.nih.gov/11794216/)]
42. Rubin GJ, Brewin CR, Greenberg N, Simpson J, Wessely S. Psychological and behavioural reactions to the bombings in London on 7 July 2005: cross sectional survey of a representative sample of Londoners. *BMJ* 2005 Sep 17;331(7517):606 [FREE Full text] [doi: [10.1136/bmj.38583.728484.3A](https://doi.org/10.1136/bmj.38583.728484.3A)] [Medline: [16126821](https://pubmed.ncbi.nlm.nih.gov/16126821/)]
43. Istepanian RS, Zhang YT. Guest editorial. Introduction to the special section: 4G Health--the long-term evolution of m-Health. *IEEE Trans Inf Technol Biomed* 2012 Jan;16(1):1-5. [doi: [10.1109/TITB.2012.2183269](https://doi.org/10.1109/TITB.2012.2183269)] [Medline: [22271836](https://pubmed.ncbi.nlm.nih.gov/22271836/)]
44. 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/)]
45. Shalev AY, Freedman S, Peri T, Brandes D, Sahar T. Predicting PTSD in trauma survivors: prospective evaluation of self-report and clinician-administered instruments. *Br J Psychiatry* 1997 Jun;170:558-564. [Medline: [9330024](https://pubmed.ncbi.nlm.nih.gov/9330024/)]
46. O'Donnell ML, Creamer MC, Parslow R, Elliott P, Holmes AC, Ellen S, et al. A predictive screening index for posttraumatic stress disorder and depression following traumatic injury. *J Consult Clin Psychol* 2008 Dec;76(6):923-932. [doi: [10.1037/a0012918](https://doi.org/10.1037/a0012918)] [Medline: [19045961](https://pubmed.ncbi.nlm.nih.gov/19045961/)]
47. Sijbrandij M, Olf M, Opmeer BC, Carlier IV, Gersons BP. Early prognostic screening for posttraumatic stress disorder with the Davidson Trauma Scale and the SPAN. *Depress Anxiety* 2008 Dec;25(12):1038-1045. [Medline: [19016483](https://pubmed.ncbi.nlm.nih.gov/19016483/)]
48. Zatzick D, Rivara F, Jurkovich G, Russo J, Trusz SG, Wang J, et al. Enhancing the population impact of collaborative care interventions: mixed method development and implementation of stepped care targeting posttraumatic stress disorder and related comorbidities after acute trauma. *Gen Hosp Psychiatry* 2011;33(2):123-134 [FREE Full text] [doi: [10.1016/j.genhosppsy.2011.01.001](https://doi.org/10.1016/j.genhosppsy.2011.01.001)] [Medline: [21596205](https://pubmed.ncbi.nlm.nih.gov/21596205/)]
49. Cucciare MA, Weingardt KR, Greene CJ, Hoffman J. Current trends in using Internet and mobile technology to support the treatment of substance use disorders. *Curr Drug Abuse Rev* 2012 Sep;5(3):172-177. [Medline: [22571448](https://pubmed.ncbi.nlm.nih.gov/22571448/)]

Abbreviations

AMC: Academic Medical Center

CAPS: Clinician Administered PTSD Scale

CBT: cognitive behavioral therapy

GP: general practitioner

HADS: Hospital Anxiety and Depression Scale

HADS-A: Hospital Anxiety and Depression Scale-Anxiety subscale

HADS-D: Hospital Anxiety and Depression Scale-Depression subscale

IES-R: Impact of Event Scale-Revised

IQR: interquartile range

LGMM: latent growth mixture modeling

MDD: major depressive disorder

MINI-Plus: Mini International Neuropsychiatric Interview

PTSD: posttraumatic stress disorder

RCT: randomized controlled trial

SF-HLQ: Short Form Health and Labour Questionnaire

TBI: traumatic brain injury

TiC-P: Trimbos/iMTA Questionnaire for Costs Associated with Psychiatric Illness

VAS: Visual Analogue Scale

VUmc: VU University Medical Center

Edited by G Eysenbach; submitted 28.11.12; peer-reviewed by N Kar, A Shalev; comments to author 28.03.13; revised version received 03.05.13; accepted 10.06.13; published 13.08.13.

Please cite as:

Mouthaan J, Sijbrandij M, de Vries GJ, Reitsma JB, van de Schoot R, Goslings JC, Luitse JSK, Bakker FC, Gersons BPR, Olf M
Internet-Based Early Intervention to Prevent Posttraumatic Stress Disorder in Injury Patients: Randomized Controlled Trial
J Med Internet Res 2013;15(8):e165

URL: <http://www.jmir.org/2013/8/e165/>

doi: [10.2196/jmir.2460](https://doi.org/10.2196/jmir.2460)

PMID: [23942480](https://pubmed.ncbi.nlm.nih.gov/23942480/)

©Joanne Mouthaan, Marit Sijbrandij, Giel-Jan de Vries, Johannes B Reitsma, Rens van de Schoot, J Carel Goslings, Jan SK Luitse, Fred C Bakker, Berthold PR Gersons, Miranda Olf. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.08.2013. 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

Participants, Usage, and Use Patterns of a Web-Based Intervention for the Prevention of Depression Within a Randomized Controlled Trial

Saskia M Kelders^{1,2}, PhD; Ernst T Bohlmeijer¹, PhD; Julia EWC Van Gemert-Pijnen³, PhD

¹Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

²National Institute for Public Health and the Environment, Bilthoven, Netherlands

³Center for eHealth Research and Disease Management, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

Corresponding Author:

Saskia M Kelders, PhD

Department of Psychology, Health and Technology

University of Twente

PO Box 217

Enschede, 7500 AE

Netherlands

Phone: 31 651070689

Fax: 31 534892388

Email: s.m.kelders@utwente.nl

Abstract

Background: Although Web-based interventions have been shown to be effective, they are not widely implemented in regular care. Nonadherence (ie, participants not following the intervention protocol) is an issue. By studying the way Web-based interventions are used and whether there are differences between adherers (ie, participants that started all 9 lessons) and nonadherers, more insight can be gained into the process of adherence.

Objective: The aims of this study were to (1) describe the characteristics of participants and investigate their relationship with adherence, (2) investigate the utilization of the different features of the intervention and possible differences between adherers and nonadherers, and (3) identify what use patterns emerge and whether there are differences between adherers and nonadherers.

Methods: Data were used from 206 participants that used the Web-based intervention Living to the full, a Web-based intervention for the prevention of depression employing both a fully automated and human-supported format. Demographic and baseline characteristics of participants were collected by using an online survey. Log data were collected within the Web-based intervention itself. Both quantitative and qualitative analyses were performed.

Results: In all, 118 participants fully adhered to the intervention (ie, started all 9 lessons). Participants with an ethnicity other than Dutch were more often adherers ($\chi^2_1=5.5$, $P=.02$), and nonadherers used the Internet more hours per day on average ($F_{1,203}=3.918$, $P=.049$). A logistic regression showed that being female (OR 2.02, 95% CI 1.01-4.04; $P=.046$) and having a higher need for cognition (OR 1.02; 95% CI 1.00-1.05; $P=.02$) increased the odds of adhering to the intervention. Overall, participants logged in an average of 4 times per lesson, but adherers logged in significantly more times per lesson than nonadherers ($F_{1,204}=20.710$; $P<.001$). For use patterns, we saw that early nonadherers seemed to use fewer sessions and spend less time than late nonadherers and adherers, and fewer sessions to complete the lesson than adherers. Furthermore, late nonadherers seemed to have a shorter total duration of sessions than adherers.

Conclusions: By using log data combined with baseline characteristics of participants, we extracted valuable lessons for redesign of this intervention and the design of Web-based interventions in general. First, although characteristics of respondents can significantly predict adherence, their predictive value is small. Second, it is important to design Web-based interventions to foster adherence and usage of all features in an intervention.

Trial Registration: Dutch Trial Register Number: NTR3007; <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3007> (Archived by WebCite at <http://www.webcitation.org/6ILhI3rd8>).

(*J Med Internet Res* 2013;15(8):e172) doi:[10.2196/jmir.2258](https://doi.org/10.2196/jmir.2258)

KEYWORDS

Web-based intervention; depression; use patterns; usage; adherence; design; engagement; attrition

Introduction

Depression has a high prevalence that poses a large burden on the health care system. Research shows that easily accessible interventions for indicated prevention (targeted at people at risk) are essential and can be cost-effective [1-3]. Web-based preventive interventions are seen as a possible format for these interventions and have been shown to be effective in reducing depressive symptoms [4-9].

Although Web-based interventions have been shown to be effective, Web-based interventions are still not widely implemented in regular care [10-13]. An issue is that not all Web-based interventions achieve the desired effects and many interventions struggle with the issue of nonadherence (ie, participants not following the intervention protocol) [10,11,14-16]. Although it is difficult to investigate a causal relationship of adherence with the effectiveness of Web-based interventions, studies have shown a relationship between adherence and increased effect of an intervention (ie, dose-effect relationship) [17,18].

In recent years, adherence has gained considerable attention. Eysenbach coined the phrase law of attrition [15], and since then there have been studies and reviews about the relationship between characteristics of participants and adherence (eg, [14,19]) and between characteristics of interventions and adherence [16,20,21]. In this study, we see adherence as following the intervention protocol (ie, using an intervention as intended by the developers), for example, completing all lessons. Although the earlier mentioned studies give insight into adherence as an outcome measure and give some recommendations how to plan for adherence, the process of adherence remains unclear. By studying the way Web-based interventions are used and whether there are differences between adherers and nonadherers, more insight can be gained into this process of adherence. Furthermore, it may be possible to extract design recommendations from this usage data and recommended use patterns for participants to increase the likelihood of adhering to the intervention.

There has been research into the usage and use patterns of Web-based interventions. Descriptive studies of freely accessible interventions have shown that they attract a considerable number of visitors, but that these visitors often interact with or access a fraction of what is possible in the intervention [22-30]. Furthermore, many studies have found that increased usage of particular features, such as completing assessments and self-monitoring, increased the effectiveness of the intervention [22,24,25,28-31]. However, insight into the way individuals use an intervention is still lacking. Particularly, insight into the patterns of use of individual participants may provide the foundation for design recommendations. Furthermore, this could lead to the formulation of usage patterns that are most likely to lead to adherence.

In addition to adherence as a process, there are still many questions regarding characteristics of respondents that may

predict adherence. Studies have investigated the predictive value of demographics and disease-related measures (eg, [14,19]), but although significant predictors have been identified, the predictive value remains low and there has been a call for investigation of other characteristics that might prove to be more predictive [10,14,16,19]. The need for cognition and the need to belong might be such characteristics. The need for cognition refers to an individual's tendency to engage in and enjoy effortful cognitive endeavors [32]. It is shown that people with a high need for cognition are more likely to engage in online activities that are more cognitively challenging [33]. As many Web-based interventions rely heavily on text and on cognitive effort to process information, it might be that individuals with a high need for cognition are more likely to adhere to a Web-based intervention. Furthermore, it has been proposed that higher levels of interactivity on health websites will lead to greater comprehension of the content, as a function of need for cognition [34], which predicts a relationship between need for cognition and adherence to Web-based interventions. The need to belong was introduced by Baumeister and Leary [35] and reflects that this desire to form interpersonal attachments is a fundamental motive that has important consequences for social functioning. Although the authors argue that the need to belong should be present to some degree in all humans in all cultures, they state that individual differences are to be expected [35]. In the context of Web-based interventions, which can be social in nature but are often something that is to be done alone, the need to belong may be a predictor for adherence (ie in Web-based interventions which are low in socialness); a higher need to belong may increase the likelihood for nonadherence.

This paper presents analyses of log data collected in a study into the adherence and effectiveness of a Web-based intervention for the prevention of depression, in which 118 of the 239 participants (49.4%) adhered to the intervention (ie, started all 9 lessons) [36]. The aims of the current study are (1) describe the characteristics of participants and investigate their relationship with adherence, (2) investigate the utilization of the different features of the intervention and possible differences between adherers and nonadherers, and (3) identify what use patterns emerge and whether there are differences between adherers and nonadherers.

Methods

Parent Study and Participants

The analyses described in this paper were performed on data collected in the parent study on the adherence and effectiveness of the Web-based intervention for the prevention of depression [36]. The parent study employed a fractional factorial experimental randomized controlled trial (RCT; NTR3007) design in which the influence of 5 components on adherence and clinical effectiveness of the Web-based intervention was studied using 8 intervention arms. This design entails that of each component, 2 levels were created and that each level of each component is present in half of the intervention arms.

Participants were adults with mild to moderate depressive symptoms (>9 and <39 on the Center of Epidemiological Studies Depression scale; CES-D [37]) who completed our online screening procedure. For the current study, data from all participants who started the first lesson were used. Therefore, we used the data from 206 of 239 participants in the parent study. Detailed information on the participants, procedures, and design of the parent study can be found in [Multimedia Appendix 1](#).

Intervention

Following Van Gemert-Pijnen et al [10], we viewed a Web-based intervention as the whole of the content, system, and the service it provides. In this conceptualization, interaction is not content, system, or service, but rather it is an integral part of a Web-based intervention and, depending on the viewpoint, it can be regarded as belonging to either category. Subsequently, we will describe the intervention, Living to the Full, according to these categories. During the study, no changes were made to the Web-based intervention apart from fixing minor bugs.

Content

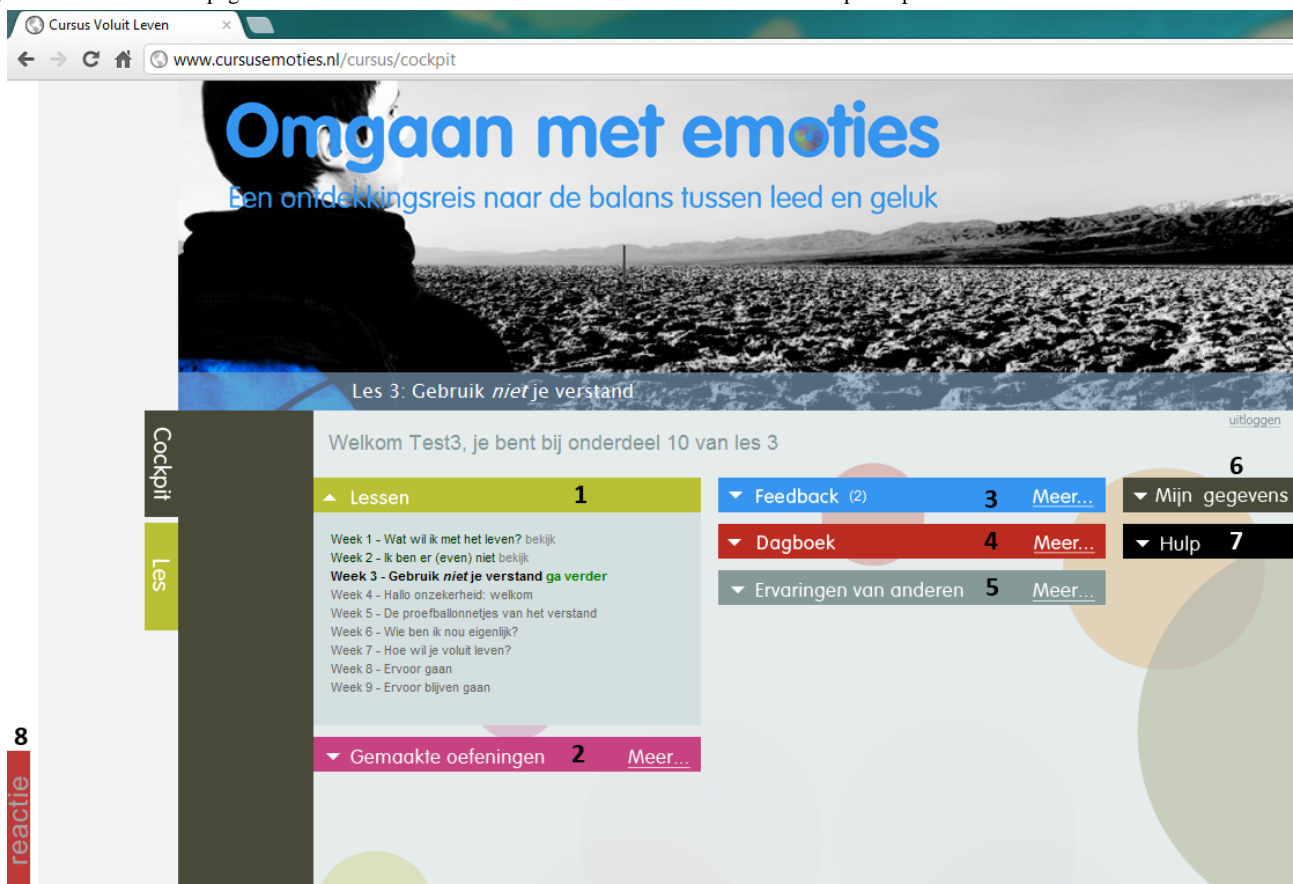
The Web-based intervention called Living to the Full is based on Acceptance and Commitment Therapy (ACT) [38] and mindfulness [39,40] and has been published as a self-help book [41]. The intervention has been shown to be effective in reducing depressive and anxiety symptoms as a group course and as a self-help course with email support [42-44]. The Web-based

intervention included 9 chronological lessons, each lesson consisting of psycho-educational material and exercises. These 9 lessons can be divided into 4 segments: part 1 (lesson 1) focuses on the view that forms the basis of the course; part 2 (lessons 2 and 3) focuses on becoming aware of coping strategies, their short term effectiveness, and lack of long term effectiveness; part 3 (lessons 4-6) focuses on learning the skills to accept suffering; and part 4 (lessons 7-9) focuses on the application of the learned lessons to daily life. The participants were asked to complete exercises both online and offline. Online exercises consisted of free-text questions, multiple-choice questions, and monitoring behavior in the Web-based diary, among others. Offline exercises consisted of practicing mindfulness, performing chosen actions, and practicing cognitive defusion, among others.

System

The intervention was developed employing methods from the CeHRes Roadmap for eHealth development [10] and this process is described elsewhere [45]. When logging on to the Web-based intervention, participants start in their cockpit (Figure 1). From there, they can access all elements of the intervention. The elements that were included for all participants were (1) lessons, (2) overview of completed exercises, (3) feedback, (4) diary, (5) success stories, (6) my account, (7) help, and (8) a “react” button which allowed respondents to comment on the application.

Figure 1. Personal home page of the Web-based intervention with the elements included for all participants.



Service

For this study, the Web-based intervention was implemented in a research setting, namely at the University of Twente, the Netherlands. Participants could access the Web-based intervention at any time, from any place, free of charge. After finishing a lesson, participants could proceed to the next lesson after receiving feedback. This feedback was provided when a participant viewed all psycho-educational material and completed all exercises. Furthermore, feedback was sent at least 5 days after the participant started the lesson (see [Multimedia Appendix 2](#) for the exact moment of feedback which differed for the levels of the support component). Participants were instructed to complete 1 lesson per week, but had 12 weeks in total to complete the 9 lessons. Participants were free to choose whether they worked through a lesson in 1 session or in multiple sessions. It was estimated that participants would spend an average of 3 hours per week on the intervention (online and offline activities combined).

Interaction

Web-based interaction with the system consisted of doing online exercises, using multimedia content, and using personalized features. Interaction in the form of feedback messages (human or automated) was provided within the system as well. Furthermore, interaction with the system occurred through automated email messages that were sent to the participants' email address to remind them to start, continue, or complete a lesson. For participants who signed up for short message service (SMS) coaching (see following paragraph), interaction also took place via their mobile phone. This interaction was 1-directional; there was no possibility to reply. Furthermore, all participants had the opportunity to contact the research staff by telephone, although this possibility was rarely used (approximately 5 phone calls during the intervention period in total).

Intervention Components

Overview

Although the components of the intervention are not the focus of this study, this section will give a short overview of each of the levels of the components to be able to place the data presented in this study in its context. A detailed description can be found in [Multimedia Appendix 2](#), and the foundations of these components can be found the parent study [36]. Each of the 8 intervention arms employed a different combination of levels of the intervention components. An overview of the composition of each of the intervention arms can be found in [Multimedia Appendix 1](#).

Support

The source of support was either human or automated. To isolate the effect of the source of support, both conditions were designed as comparable as possible regarding length of feedback messages, tailored content, and presentation (including a photo of the counselor). To maintain the unique differences between human and automated support (increased possibility for interaction in human support and the increased possibility for timely feedback in automated support), participants in the human support condition had the opportunity to ask questions to their counselor, and participants in the automated support condition

received 1 additional Web-based instant feedback message per lesson.

SMS Coaching

Participants in the condition that included SMS text messages had the opportunity to turn the SMS coach on. This SMS coach sent 3 predesigned text messages each week to a mobile phone number provided by the participant. The text messages were written by the researchers before the study started and the content was based on the results of the development study of the intervention [45]. Each week, 3 SMS text messages were sent containing motivational, mindfulness, and content-related information. All SMS text messages were presented in the text message tab of the application, independent of whether the SMS coach was turned on or off, but only for the participants in the condition that included text messages.

Experience Through Technology

The high experience condition contained additional multimedia and interactive material in the form of short movies, interactive exercises, and multimedia presentations of metaphors.

Tailoring of Success Stories

The intervention contained a success story for each lesson. For the high-tailored condition, each success story was tailored on 4 of the following aspects: gender, age, marital status, daily activity, most prominent symptom, and reason for participating in the Web-based intervention. The stories were tailored to a different combination of aspects each week and not on all aspects to maintain the credibility of the stories. In the low-tailored condition, a standard success story was presented each week.

Personalization

The high-personalization condition included personalized content that was adapted (the system shows the motto and picture selected by the participant; the system shows the most important values selected by the participant) and adaptable (possibility to create a personal top 5 of aspects from the course that the participant found most important).

Data Collection and Analysis

Characteristics of participants were collected at baseline by using an online questionnaire. Depressive symptoms were measured with the CES-D (20 items, score 0-60; higher=more depressive symptoms [37,46]), anxiety symptoms were measured with the Hospital Anxiety and Depression Scale (HADS-A; 7 items, score 0-21; higher scores=more anxiety symptoms [47,48]). Need for cognition was measured using the Need for Cognition Short Form (18 items, score -54-54; higher scores=more need for cognition [32]). Need to belong was measured using the Need to Belong Scale (10 items, mean score 1-7; higher scores=more need to belong [49]). Internet usage was measured using 1 item (ie, "On average, for how many hours do you use the Internet per day?"). Internet experience was measured by using 10 items of the following format: "Do you ever use the following Internet applications?" The 10 items focused on the usage of search engines, webmail, online shopping, online banking, online communities, photo and video websites, (micro)blogs, chat, radio or music websites, and online

(health) courses. The score was attained by counting the number of items that were answered with at least once in a while (range 0-10).

Usage of the Web-based intervention was measured objectively by log files. From these log files, adherence could be extracted. Adherence was defined as a participant starting lesson 9, because the intervention is intended to be used during the 9 lessons.

The log files contained a record of actions taken by each participant with for each action the following information: unique participant identification number, action type, action specification, and time and day. The action types that were logged were log-in, log-out, start lesson, start mindfulness, download mindfulness, view success story, view feedback message, start video, turn on SMS coach, turn off SMS coach, and view text message. Action specifications were, for example, the name of the mindfulness exercise started or which text message was viewed.

Descriptive analyses of use patterns were performed on 20 arbitrarily selected participants; 5 early nonadherers (ie, did not start lesson 5), 5 late nonadherers (ie, started lesson 5 but did not start lesson 9), and 10 adherers. We divided the nonadherers into early and late nonadherers to explore whether there were differences between these groups. It may be that people who nonadhere early have different reasons for nonadhering than late nonadherers. These early reasons may be more general aspects that become clear at an early stage (eg, the content is not attractive to them or the format of the intervention does not match their expectations). Late reasons may be more related to the process of the intervention or to the motivation (eg, it is hard to spend enough time each week on the intervention). It may be that late nonadherers are more similar to adherers and are easier to persuade to become adherers, whereas for early nonadherers, the intervention may simply not be suitable. Although the reasons for early or late nonadherence cannot be derived from this research, the results can show whether late nonadherers are more similar to adherers regarding their usage of the intervention. Effort was made to ensure that selected participants had the same distribution of demographic characteristics and randomized group as the full sample. Furthermore, we only selected participants who did not start to nonadhere in lessons 2, 5, or 8 because these were the lessons we investigated and we wanted to avoid including patterns of participants who did not complete the lessons under investigation. See [Multimedia Appendix 3](#) for an overview of demographics, randomized group, and lesson reached of these selected participants. Of these participants, we examined all actions in lesson 2 (all selected participants), lesson 5 (late nonadherers and adherers), and lesson 8 (adherers only) to identify emerging use patterns. We chose to examine these lessons because they reflect the 3 main segments of the content of the intervention and because we wanted to avoid the first and the last lesson for the expected nonregular use pattern in these lessons; we expect the participants to explore and get to know the application more in the first lesson and the last lesson is shorter (ie, less text and exercises) than the other lessons. Of each lesson and for each selected participant, we recorded all actions between the time they started the lesson under

investigation and the time they started the following lesson. Moreover, the number of sessions (a log-in and following actions until a log-out action or a period of 30 minutes of inactivity was counted as 1 session) was derived, as well as the total duration of these sessions and the time between session. Furthermore, the number of sessions used to complete all exercises and content of the lesson were counted. We chose to do this analysis only for a small subsample of the data because the focus of this exploratory analysis was on pattern recognition related to use of features of the interventions. Furthermore, the choice was pragmatic because of the lack of software to analyze log files, all analyses were done by hand.

Statistical analyses were done using PASW 18 (Predictive Analytics Software; IBM, USA). Differences between adherers and nonadherers were investigated using 1-way analyses of variance (ANOVA) and chi-square tests (χ^2). Logistic regression was used to assess whether baseline characteristics predicted adherence. Because of the exploratory nature of the logistic regression, all predictor variables were added at once, using the enter method.

Results

Participant Characteristics

Baseline demographics and outcome measures of the 206 participants who used the intervention are shown in [Table 1](#). There were differences between adherers and nonadherers on ethnicity (participants with an ethnicity other than Dutch were more often adherers; $\chi^2_1=5.5$, $P=.02$) and Internet usage (nonadherers used the Internet significantly more hours per day on average; $F_{1,203}=3.918$, $P=.049$). Women were more often adherers, but this did not meet statistical significance ($\chi^2_1=3.7$, $P=.05$). Nonadherers had a higher need to belong, but this did not meet statistical significance ($F_{1,204}=3.133$, $P=.08$).

Adherence

The average number of lessons started was 6.9 out of a possible 9, and 57% of the participants in this study completely adhered to the intervention (mode and median = 9 lessons). [Figure 2](#) shows the percentage of participants who reached a certain lesson. From this figure, the largest group of nonadherers began to nonadhere in lesson 2, followed by lessons 3 and 6. Moreover, we can see that 26.2% (54/208) of participants were early nonadherers (ie, did not start lesson 5) and 16.6% (34/208) were late nonadherers (ie, started lesson 5 but did not start lesson 9).

To explore the possible predictive value of baseline characteristics for adherence (ie, starting all 9 lessons), we performed an exploratory logistic regression with all baseline characteristics showed in [Table 1](#) entered as predictors. [Table 2](#) shows that significant predictors in the model were gender and need for cognition, in which being female and having a higher need for cognition increased the odds of adhering to the intervention. A linear regression to predict the lesson reached by baseline characteristics yielded a significant model ($\chi^2_{12}=28.9$, $P=.004$; Cox & Snell $R^2=0.132$, Nagelkerke $R^2=0.177$), but no significant predictor variables (data not shown).

Table 1. Baseline demographics and outcome measures of all participants, and differences between adherers and nonadherers.

Participant characteristic	Total (N=206)	Adherers (n=118)	Nonadherers (n=88)	<i>P</i>
Age (years), mean (SD)	44.7 (12.5)	45.2 (12.6)	43.9 (12.3)	.47
Gender (women), n (%)	150 (72.8)	92 (78.0)	58 (65.9)	.05
Ethnicity, n (%)				.02
Dutch	188 (91.3)	103 (87.3)	85 (96.6)	
Other	18 (8.7)	15 (12.7)	3 (3.4)	
Education, n (%)				.51
High	139 (67.5)	82 (69.5)	57 (64.8)	
Middle	53 (25.7)	30 (25.4)	23 (26.1)	
Low	14 (6.8)	6 (5.1)	8 (9.1)	
Marital status, n (%)				.46
Married	72 (35.0)	45 (38.1)	27 (30.7)	
Divorced	41 (19.9)	20 (16.9)	21 (23.9)	
Widowed	4 (1.9)	3 (2.5)	1 (1.1)	
Unmarried	89 (43.2)	50 (42.4)	39 (44.3)	
Daily activities, n (%)				.15
Paid job	131 (63.6)	69 (58.5)	62 (70.5)	
Student	16 (7.8)	9 (7.6)	7 (8.0)	
No job	59 (28.6)	40 (33.9)	19 (21.6)	
CES-D, mean (SD)	24.9 (6.9)	24.5 (7.3)	25.4 (6.5)	.35
HADS-A , mean (SD)	9.7 (2.6)	9.4 (2.5)	10.0 (2.6)	.13

Table 2. Logistic regression baseline characteristics and adherence.

Included	B ^a (SE)	<i>P</i>	OR (95% CI)
Constant	0.56 (1.82)	.76	
Age	-0.01 (0.02)	.65	0.99 (0.96-1.02)
Gender	0.70 (0.35)	.046	2.02 (1.01-4.04)
Ethnicity	1.29 (0.70)	.07	3.63 (0.92-14.26)
Education	0.30 (0.26)	.25	1.35 (0.81-2.24)
Marital status	-0.09 (0.14)	.53	0.92 (0.70-1.20)
Daily activities	0.35 (0.19)	.08	1.41 (0.97-2.06)
CES-D	-0.01 (0.02)	.71	0.99 (0.95-1.04)
HADS-A	-0.12 (0.07)	.07	0.89 (0.78-1.01)
Need for Cognition	0.02 (0.01)	.02	1.02 (1.00-1.05)
Need to Belong	-0.33 (0.27)	.22	0.72 (0.43-1.21)
Internet usage	-0.16 (0.09)	.06	0.85 (0.72-1.01)
Internet experience	-0.05 (0.11)	.64	0.95 (0.77-1.18)

^aB: unstandardized coefficient.

Figure 2. Graph of lessons completed against proportion of participants.

Usage

From the log files, the number of times each participant performed an action in the Web-based application was extracted (Table 3). Overall, participants logged in an average of 4 times per lesson, but adherers logged in significantly more times per lesson started than nonadherers ($F_{1,204}=20.710$; $P<.001$). Other differences were that adherers downloaded a higher percentage of possible unique mindfulness exercises than nonadherers ($F_{1,204}=5.888$; $P=.02$) and that adherers in the condition that included SMS coaching viewed a larger percentage of the possible text messages than nonadherers in that condition ($F_{1,103}=7.668$; $P=.007$). To explore whether intervention components influenced the frequency of user actions, we compared the percentage of unique success stories that were viewed between participants in the condition with high- and low-tailored success stories and found that there was no significant difference. However, there was a difference between the total number of unique feedback messages viewed between the conditions with human and automated support (whole group: human support 10.7 unique messages viewed; automated support 5.9 unique messages viewed; $F_{1,204}=37.322$, $P<.001$) and between the conditions on the number of unique messages viewed per lesson for adherers as well as for nonadherers (adherers: human support 1.7 per lesson, automated support 0.9 per lesson, $F_{1,116}=93.604$, $P<.001$; nonadherers: human support 1.1 per lesson, automated support 0.6 per lesson, $F_{1,86}=23.860$, $P<.001$).

Use Patterns

To examine in more detail the way participants interacted with the system during the lessons, the use patterns of 20 participants

(5 early nonadherers, 5 late nonadherers, and 10 adherers) on lesson 2 (all selected participants), lesson 5 (late nonadherers and adherers), and lesson 8 (adherers only) were investigated. Multimedia Appendix 4 presents all actions per participant per lesson, organized into sessions. Furthermore, Multimedia Appendix 4 presents the duration of each session, the time in between sessions, and an overview of the total duration of sessions and time between sessions per participant per lesson. A summary of this information for early nonadherers, late nonadherers, and adherers is presented in Table 4. From this table we can see that there seem to be differences between the use patterns of the 3 groups. First, early nonadherers used less sessions and spent less time than late nonadherers and adherers, and used less sessions to complete the lesson than adherers. Second, late nonadherers had a shorter total duration of the sessions than adherers, with the difference being more pronounced in lesson 5. Finally, adherers used less sessions (total and to complete a lesson) in the later lessons, but there was no visible trend for the duration of sessions and time between sessions, although they were a higher for lesson 5. When looking at the data in Multimedia Appendix 4, we observed some notable patterns:

1. There are many sessions that involve only a log-in and a log-out action, with less than a minute in between.
2. Adherers start the later lessons with a very short first session.
3. Many feedback messages are not read the first session after they are available.
4. There are many log-in actions shortly after another action.

Table 3. User actions of adherers and nonadherers.

User actions	Adherers (n=118)	Nonadherers (n=88)	Total (N=206)
Log-in, ^a mean (SD)			
Total	40.2 (19.8)	14.4 (13.6)	29.1 (21.6)
Per lesson	4.5 (2.2)	3.2 (1.5)	3.9 (2.0)
Feedback messages viewed, mean (SD)			
Total	22.9 (17.6)	6.1 (7.8)	15.7 (16.5)
Unique messages	12.0 (5.2)	3.8 (3.7)	8.5 (6.1)
Unique messages per lesson	1.3 (0.6)	0.8 (0.6)	1.1 (0.6)
Mindfulness exercises			
Total started, mean (SD)	7.8 (5.6)	3.6 (3.3)	6.0 (5.2)
Unique started, mean (SD) % ^b	3.6 (1.4) 72.0%	2.1 (1.3) 74.3%	2.9 (1.6) 73.0%
Unique downloaded, mean (SD) % ^b	2.6 (2.1) 51.5%	1.1 (1.3) 37.7%	1.9 (1.9) 45.6%
Unique used, mean (SD) % ^b	4.4 (1.0) 87.6%	2.3 (1.3) 81.6%	3.5 (1.5) 85.0%
Success stories viewed			
Total, mean (SD)	8.8 (7.5)	3.5 (3.5)	6.5 (6.7)
Unique, mean (SD) % ^b	5.2 (2.8) 57.3%	2.4 (1.9) 61.4%	4.0 (2.8) 59.1%
Text message coaching ^c			
Participants that turned text message coaching on, n ^d	19	7	26
Lessons turned on, mean (SD) ^e	7.9 (2.6)	2.4 (1.7)	6.5 (3.4)
Total messages viewed, mean (SD)	14.3 (20.0)	2.4 (3.7)	9.6 (16.7)
Unique messages viewed, mean (SD) % ^b	8.4 (8.9) 31.0%	1.8 (2.8) 14.9%	5.8 (7.8) 24.6%
Movies viewed ^f			
Total, mean (SD)	5.4 (6.1)	2.0 (3.8)	3.9 (5.5)
Unique, mean (SD) % ^b	3.5 (3.4) 38.6%	1.3 (2.3) 25.5%	2.5 (3.2) 32.9%

^aLog-ins within 30 minutes of the previous log-in were not counted to make the log-ins reflect the number of sessions more accurately; Log-ins per started lesson: number of log-ins divided by the number of the last lesson started.

^b% = unique actions/possible actions. For adherers, the number of possible actions is the total number of available actions of that kind in the whole intervention. For nonadherers, the number of possible actions is the total number of available actions in all lessons that the participant started.

^cOnly for participants in the condition that included text message coaching; n=105; adherers n=63; nonadherers n=42.

^dThe number of participants that turned the text message coach on at least 1 time.

^eThe number of lessons the text message coach was turned on for the participants that turned the text message coach on at least 1 time.

^fOnly for participants in the high experience condition; n=116; adherers n=65; nonadherers n=51.

Table 4. Mean number of sessions and duration for early nonadherers (n=5), late nonadherers (n=5), and adherers (n=10).

Variable	Early nonadherers, mean (SD)	Late nonadherers, mean (SD)	Adherers, mean (SD)
Lesson	2	2	5
Total sessions	2.8 (1.6)	4.4 (1.5)	4.0 (1.6)
Sessions to complete lesson	1.8 (0.8)	2.0 (1.2)	2.8 (1.6)
Total duration of session (min)	36.2 (44.8)	64.0 (45.2)	38.8 (33.3)
Time in between sessions (days)	6.7 (4.1)	10.0 (4.1)	10.8 (1.8)

Discussion

Principal Results

The aims of this study were to (1) describe the characteristics of participants and investigate their relationship with adherence, (2) investigate the utilization of the different features of the intervention and possible differences between adherers and nonadherers, and (3) identify what use patterns emerge and whether there are differences between adherers and nonadherers.

The participants in this study were primarily Dutch females with a higher education level and a paid job. This group is similar to the group reached by many Web-based or eHealth interventions (eg, [15,19,22]), and this was the expected group which we took into account in the development process. When looking at differences between adherers and nonadherers, we see that although we reached only a small percentage of participants with an ethnicity other than Dutch; these participants were more often adherers. Others have stressed the importance and challenge of reaching people with non-Dutch ethnicity [13]. This study shows that if we can succeed in reaching this population, it may be easier to keep them engaged with a Web-based intervention, but this needs further research. Furthermore, nonadherers generally used the Internet for more hours per day than adherers. This finding is similar to other studies [19,50] and is something that deserves more research. One possible explanation is that people who differ in their amount of Internet use, also differ in the expectations they have of Web-based systems and, in this case, Web-based interventions. It may be that this Web-based intervention does not completely fit the mental model of a Web application of regular Internet users; the Web-based intervention, for example, may require more intense use as opposed to browsing where information is screened and many pages are viewed in a short amount of time. Our logistic regression model to predict adherence from characteristics of participants had relatively low predictive power (Nagelkerke $R^2=0.177$) in which only being female and having a higher need for cognition increased the odds of adhering to the intervention. The finding that women are more likely to adhere was mirrored in the finding that more women were adherers versus nonadherers (although this was statistically nonsignificant) and may reflect our choice to include more women as participants in the development process. Moreover, it strengthens the assumption that it is important to take the target group into account. If we intend to reach and engage men more, we should redesign the intervention using their input. The second significant predictor was the need for cognition, which supports our hypothesis that a higher need for cognition may be beneficiary for completing a Web-based intervention that relies substantially on text and on cognitive effort to process information. This implies that if an intervention is not only aimed at participants with a high need for cognition, attention should be paid to make the intervention more suitable for participants with a lower need for cognition. Although this model and other studies [14,19,51] show that individual differences play a role in adherence, the predictive value of the characteristics we measured is still small. A different approach has been used in the field of persuasive technology, where tailoring persuasive messages to personality traits has been

shown to be effective [52]. Furthermore, in this field the question why certain individuals are persuaded and others are not has been investigated from a more practical view: by generating an individual persuasion profile from data on actual behavior, the most effective strategy to persuade this individual can be deduced and employed [53]. From there, one can theorize where these persuasion profiles come from and whether they can be measured in advance. This might be a practical way to tackle this issue in the field of Web-based interventions and eHealth.

Overall, of the 206 participants that used the application, 118 participants adhered to the intervention. Although we included the percentage of adherers by using these numbers, it should be noted that we only report on participants that started lesson 1. The true adherence derived from all participants is 49.4% (118/239) [36]. This percentage is in-line with the average adherence found in a systematic review [16]. The data showed that most of the participants who did not adhere to the intervention, started to nonadhere during the first 3 lessons (55%, 48/88 participants that started the first lesson, but did not adhere to the intervention). This might reflect the content of the intervention, in which the first lesson focuses on whether the participants are really open for the therapy and the next 2 lessons focus on becoming aware that the coping strategies they use are not effective. This can be very confronting and may, therefore, explain the high nonadherence in these lessons. Interestingly, there is also a fairly large group of participants that start to nonadhere during lesson 6. This lesson is the last lesson in the segment on learning new skills to accept suffering, and this particular lesson focuses on the observing self. Counselors who have given this course know that this is hard lesson for many participants, which may explain the large group of nonadherers in this lesson. For the redesign of this Web-based intervention, this finding indicates that this might be a moment when the intervention should provide extra motivation, for example, through more interaction or simply by acknowledging that it is known that this is a hard moment to stick with the program. According to the supportive accountability model, human support may increase adherence at these times by providing the right amount (tailored to the individual) of support [54].

Our results on the usage of the different features mirror the results of studies on the usage of freely available Web-based interventions in that participants do not use all the features that they can use [22-30]. It seems that features that are an integral part of the therapy (eg, the mindfulness exercises in this study) are used more than additional features (success stories, text message coaching, and movies). This is something to keep in mind when designing or redesigning Web-based interventions: be aware that not all features in an application will be used and try to integrate features into the intervention instead of adding them onto the intervention. The success stories, for example, could be integrated more into the intervention by inserting them into each lesson. This is in-line with a recommendation from Krukowski et al [30] to encourage or more prominently feature certain features to increase usage and, ideally, lead to improved outcomes.

Significant differences between the user actions of adherers and nonadherers (ie, adherers showed more log-ons per lesson,

downloaded more mindfulness exercises, and viewed more text messages than nonadherers) indicate that adherers not only have more endurance regarding usage during the full duration of the intervention, but are also more engaged with the intervention compared to nonadherers. This mirrors the finding that adherers show more involvement with the intervention [36]. This higher engagement within a lesson may be beneficial for the effectiveness of the intervention because more exposure to a Web-based intervention has been shown to increase the effectiveness (eg, [17,30]). For this specific intervention, completing a lesson in more than 1 log-in was recommended because the content requires repeated practice (eg, mindfulness and diary exercises) and time to reflect on the content before completing exercises. However, a participant could complete a full lesson in 1 log-in.

The average number of feedback messages viewed per lesson was below 1 for adherers and nonadherers, which shows that not all feedback messages have been viewed. Receiving feedback was the most wanted and expected feature of a Web-based intervention according to the participants in our development study, and providing support has been shown to have a positive effect on the effectiveness of Web-based interventions [4]. This makes the finding that not all feedback messages have been viewed even more striking. An explanation may be that this feature that was thought to be integral to the treatment by the developers was implemented in a way that did not reflect this integral nature; feedback messages were presented in a different section of the system than the lessons (the main part of the therapy) and participants could proceed to the next lesson without viewing the feedback message. Additionally, participants in the human support condition viewed more feedback messages than participants in the automated support condition. This finding is not surprising because the automated support condition included only 1 message per lesson, whereas the human support condition included the possibility to ask questions and request more feedback. Interestingly, the study into adherence and effectiveness of this intervention [36] did not show a significant difference in effectiveness at follow-up between these conditions, even though a counselor gave the feedback and more feedback messages were given (as shown in this study). These 2 findings show a need to further investigate the role of support and feedback in Web-based interventions.

Our analyses of the use patterns of 20 participants over 3 different lessons, provided us with useful insights. This more qualitative analysis confirmed our quantitative results on user actions: adherers are overall more engaged, they use more sessions, and spend more time with the intervention. Moreover, the analyses of the use patterns show us that there may be a difference between early nonadherers, late nonadherers, and adherers, in which late nonadherers are more similar to adherers in the number of sessions, but have a shorter duration of sessions, which is more similar to early nonadherers. This seems to fit our hypothesis that there is a difference between early and late reasons for nonadherence. Late nonadherers may be more similar to adherers and they may be easier to persuade to become adherers, whereas the intervention may simply not be suitable

for early nonadherers. This should be investigated in future research.

By identifying differences between adherers and nonadherers, it becomes possible to screen for these wrong patterns and identify participants that are at risk to become nonadherers. This provides the opportunity to intervene, for example, by notifying these participants that they have a use pattern that increases the likelihood for nonadherence and suggesting a more appropriate use pattern. This combination of monitoring and self-monitoring of behavior and providing suggestions for different behavior are thought to be persuasive strategies for behavior change [55,56]. A different way to intervene may be to provide the participants with more or different interaction to increase adherence and effectiveness [54]. Although the current research provides a way to intervene, the specific action that is needed at that time for a specific participant should be the focus of future research.

Our in-depth analyses of the use patterns presented in [Multimedia Appendix 4](#), yielded notable patterns that are useful for the redesign of this specific intervention. For example, the frequent log-in–log-out actions with no user action in between might be behavior of participants who were waiting for feedback. This hypothesis is supported by the finding that this pattern often occurs after participants have completed a lesson, but have not received feedback. A redesign option is to provide a prominent feature “when will I get my feedback?” where a timer can be shown with the expected time of feedback. This feature can then also be used to direct the participants to the features that they have not used at that time, to support participants to employ all the features to benefit most from the intervention. We saw that many adherers start the later lessons with a very short first session. This reflects the set-up of the intervention, in which the next lesson is only available after the participants complete the current lesson and a certain time since the start of the lesson has passed. This timer is started as soon as the lesson is started, so this first short session might be done to start the timer. The finding that many feedback messages are not read the first time they become available reflects the earlier finding that not all feedback messages are being read and might be improved by making it clearer that there is a new feedback message. A known bug in the application that has not been fixed is that a participant is logged out of the application when using the back button of the browser. This bug is a likely explanation of the many log-in actions shortly after another.

For this study, we used the log data of the Web-based intervention itself. This allowed us to identify actions of specific participants and relate them to whether the participant adhered to the intervention or not. Other studies have advocated the use of Google analytics, for example [57], but although this provides valuable information on a general level, it is not possible to identify specific participants, which diminishes the value of those methods for Web-based interventions that are intended to be used on multiple occasions. However, when developing a Web-based intervention, it is important to specify which information is important to be logged. For example, in our study, sessions were not logged as such, which meant that this had to be done manually, which is a tedious exercise. Furthermore, we manually wrote out all sessions for the selected participants

in the selected lessons. Although this method provided valuable information, it is not feasible to do this for all participants for all lessons, which entails that analyses are done on a subset of the data. More advanced methods are needed to make use of all information that is collected. One such approach might be found in the use of Markov chains as used by Tian et al [58], although this might be less feasible for Web-based interventions that are intended to be used on more occasions. Another approach might be to employ pattern recognition methods from a machine learning perspective to see whether there are different patterns for adherers and nonadherers that can be automatically recognized or learned.

Limitations

A limitation of this study is that we analyzed and interpreted log data without actively involving the participants. We did not ask participants why they used the intervention the way they used it. This information may have made it easier to interpret the data and to check whether our interpretation is correct. On the other hand, it is important to use objective log data and not to rely on subjective measures of how participants state that they used the intervention, because subjective data on usage are likely to be less accurate. Another limitation is the issue of generalizability. Our study used data from 1 intervention for the prevention of depression, which was used by primarily higher-educated Dutch women. Furthermore, we only investigated the use patterns of a small sample of these participants. The observed use patterns may be specific for this group using this intervention. However, many interventions, especially mental health interventions, have similar characteristics [16] and reach the same audience as stated earlier. Furthermore, the implications regarding designing for adherence, the limited predictive value of regular participant characteristics for adherence and the possibility to intervene based on screening

of use patterns, seem to be broader than only for this intervention with this audience.

Future Research

An interesting area for research can be found in a new way of analyzing the use patterns and investigating whether it is useful and feasible to intervene during the use of the intervention on the basis of the analyses of real-time use patterns. An earlier step might be to identify use patterns that are related to adherence and to design or redesign interventions in such a way to promote these use patterns. A different area of future research lies in the investigation of a more pragmatic way to identify participant characteristics that may influence or predict adherence, following the persuasion profiling approach [53]. Furthermore, our results indicate that the different content of lessons may need a different amount or mode of interaction. Here lies an interesting line of research: how can the content, system, service, and interaction of a Web-based intervention be attuned to one other to achieve the best match?

Conclusion

In conclusion, we can say that using log data combined with baseline characteristics of participants of the intervention Living to the Full provided valuable lessons for redesign of this intervention and the design of Web-based interventions in general. First, although characteristics of respondents can significantly predict adherence, their predictive value is small. Therefore, we should look into other ways of classifying participants to make useful predictions about how individual difference may influence adherence. Second, it is important to design Web-based interventions to foster adherence and usage of all features in an intervention. A possibility for this is a smarter system that logs the current use pattern of a participant and intervenes when necessary, for example, by providing feedback or links to features that have not been accessed yet.

Conflicts of Interest

The authors were involved in the development of the Web-based intervention Living to the Full.

Multimedia Appendix 1

Description of parent study.

[PDF File (Adobe PDF File), 124KB - [jmir_v15i8e172_app1.pdf](#)]

Multimedia Appendix 2

Detailed description of the five intervention factors.

[PDF File (Adobe PDF File), 1MB - [jmir_v15i8e172_app2.pdf](#)]

Multimedia Appendix 3

Characteristics of respondents for analyses of usage patterns.

[PDF File (Adobe PDF File), 86KB - [jmir_v15i8e172_app3.pdf](#)]

Multimedia Appendix 4

User actions, duration and time between sessions per participant per lesson.

[PDF File (Adobe PDF File), 266KB - [jmir_v15i8e172_app4.pdf](#)]

Multimedia Appendix 5

CONSORT-EHEALTH checklist V1.6.2 [59].

[PDF File (Adobe PDF File), 985KB - [jmir_v15i8e172_app5.pdf](#)]

References

1. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry* 2004 Jun;184:526-533 [FREE Full text] [Medline: [15172947](#)]
2. Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. *Am J Psychiatry* 2008 Oct;165(10):1272-1280. [doi: [10.1176/appi.ajp.2008.07091422](#)] [Medline: [18765483](#)]
3. Hollon SD, Muñoz RF, Barlow DH, Beardslee WR, Bell CC, Bernal G, et al. Psychosocial intervention development for the prevention and treatment of depression: promoting innovation and increasing access. *Biol Psychiatry* 2002 Sep 15;52(6):610-630. [Medline: [12361671](#)]
4. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009 Dec;38(4):196-205. [doi: [10.1080/16506070903318960](#)] [Medline: [20183695](#)]
5. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](#)] [Medline: [20967242](#)]
6. 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](#)]
7. Beekman AT, Smit F, Stek ML, Reynolds CF, Cuijpers PC. Preventing depression in high-risk groups. *Curr Opin Psychiatry* 2010 Jan;23(1):8-11. [doi: [10.1097/YCO.0b013e328333e17f](#)] [Medline: [19901835](#)]
8. Cuijpers P, van Straten A, Warmerdam L, van Rooy MJ. Recruiting participants for interventions to prevent the onset of depressive disorders: possible ways to increase participation rates. *BMC Health Serv Res* 2010;10:181 [FREE Full text] [doi: [10.1186/1472-6963-10-181](#)] [Medline: [20579332](#)]
9. Van Voorhees BW, Mahoney N, Mazo R, Barrera AZ, Siemer CP, Gladstone TR, et al. Internet-based depression prevention over the life course: a call for behavioral vaccines. *Psychiatr Clin North Am* 2011 Mar;34(1):167-183 [FREE Full text] [doi: [10.1016/j.psc.2010.11.002](#)] [Medline: [21333846](#)]
10. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](#)] [Medline: [22155738](#)]
11. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. *PLoS Med* 2011;8(1):e1000387 [FREE Full text] [doi: [10.1371/journal.pmed.1000387](#)] [Medline: [21267058](#)]
12. Christensen H, Reynolds J, Griffiths KM. The use of e-health applications for anxiety and depression in young people: challenges and solutions. *Early Interv Psychiatry* 2011 Feb;5 Suppl 1:58-62. [doi: [10.1111/j.1751-7893.2010.00242.x](#)] [Medline: [21208393](#)]
13. Riper H, Smit F, Van der Zanden R, Conijn B, Kramer J, Mutsaers K. E-mental health. High tech, high touch, high trust [Programmeringsstudie e-mental health in opdracht van het ministerie van VWS]. Utrecht, the Netherlands: Trimbos Institute; 2007. URL: <https://www.infopuntveiligheid.nl/Infopuntdocumenten/ICOM%20-%20Programmeringsstudie%20E-mental%20Health.pdf> [accessed 2013-08-05] [WebCite Cache ID 61e7luQjr]
14. 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](#)] [Medline: [19403466](#)]
15. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](#)] [Medline: [15829473](#)]
16. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to Web-based interventions. *J Med Internet Res* 2012;14(6):e152 [FREE Full text] [doi: [10.2196/jmir.2104](#)] [Medline: [23151820](#)]
17. 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](#)] [Medline: [21821503](#)]
18. Manwaring JL, Bryson SW, Goldschmidt AB, Winzelberg AJ, Luce KH, Cunning D, et al. Do adherence variables predict outcome in an online program for the prevention of eating disorders? *J Consult Clin Psychol* 2008 Apr;76(2):341-346. [doi: [10.1037/0022-006X.76.2.341](#)] [Medline: [18377129](#)]

19. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. *J Med Internet Res* 2011;13(2):e32 [FREE Full text] [doi: [10.2196/jmir.1624](https://doi.org/10.2196/jmir.1624)] [Medline: [21493191](https://pubmed.ncbi.nlm.nih.gov/21493191/)]
20. 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;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/)]
21. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. *Comput Inform Nurs* 2011 Feb;29(2 Suppl):TC9-T20. [doi: [10.1097/NCN.0b013e3182155274](https://doi.org/10.1097/NCN.0b013e3182155274)] [Medline: [21372645](https://pubmed.ncbi.nlm.nih.gov/21372645/)]
22. Balmford J, Borland R, Benda P. Patterns of use of an automated interactive personalized coaching program for smoking cessation. *J Med Internet Res* 2008;10(5):e54 [FREE Full text] [doi: [10.2196/jmir.1016](https://doi.org/10.2196/jmir.1016)] [Medline: [19097975](https://pubmed.ncbi.nlm.nih.gov/19097975/)]
23. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. *J Med Internet Res* 2010;12(1):e9 [FREE Full text] [doi: [10.2196/jmir.1347](https://doi.org/10.2196/jmir.1347)] [Medline: [20350926](https://pubmed.ncbi.nlm.nih.gov/20350926/)]
24. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. *J Med Internet Res* 2002 Mar;4(1):e3 [FREE Full text] [doi: [10.2196/jmir.4.1.e3](https://doi.org/10.2196/jmir.4.1.e3)] [Medline: [11956035](https://pubmed.ncbi.nlm.nih.gov/11956035/)]
25. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. *J Med Internet Res* 2010 Dec;12(4):e52 [FREE Full text] [doi: [10.2196/jmir.1430](https://doi.org/10.2196/jmir.1430)] [Medline: [21087922](https://pubmed.ncbi.nlm.nih.gov/21087922/)]
26. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. *J Med Internet Res* 2012;14(1):e4 [FREE Full text] [doi: [10.2196/jmir.1792](https://doi.org/10.2196/jmir.1792)] [Medline: [22260810](https://pubmed.ncbi.nlm.nih.gov/22260810/)]
27. 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/)]
28. Linke S, Murray E, Butler C, Wallace P. Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. *J Med Internet Res* 2007;9(2):e10 [FREE Full text] [doi: [10.2196/jmir.9.2.e10](https://doi.org/10.2196/jmir.9.2.e10)] [Medline: [17513281](https://pubmed.ncbi.nlm.nih.gov/17513281/)]
29. Zbikowski SM, Hapgood J, Smucker Barnwell S, McAfee T. Phone and Web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. *J Med Internet Res* 2008;10(5):e41 [FREE Full text] [doi: [10.2196/jmir.999](https://doi.org/10.2196/jmir.999)] [Medline: [19017583](https://pubmed.ncbi.nlm.nih.gov/19017583/)]
30. Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-based weight control: the relationship between web features and weight loss. *Telemed J E Health* 2008 Oct;14(8):775-782 [FREE Full text] [doi: [10.1089/tmj.2007.0132](https://doi.org/10.1089/tmj.2007.0132)] [Medline: [18954247](https://pubmed.ncbi.nlm.nih.gov/18954247/)]
31. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. *J Med Internet Res* 2011;13(1):e9 [FREE Full text] [doi: [10.2196/jmir.1391](https://doi.org/10.2196/jmir.1391)] [Medline: [21371992](https://pubmed.ncbi.nlm.nih.gov/21371992/)]
32. Cacioppo JT, Petty RE, Kao CF. The efficient assessment of need for cognition. *J Pers Assess* 1984 Jun;48(3):306-307. [doi: [10.1207/s15327752jpa4803_13](https://doi.org/10.1207/s15327752jpa4803_13)] [Medline: [16367530](https://pubmed.ncbi.nlm.nih.gov/16367530/)]
33. Tuten TL, Bosnjak M. Understanding differences in web usage: the role of need for cognition and the five factor model of personality. *Soc Behav Pers* 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)]
34. Lustria MLA. Can interactivity make a difference? Effects of interactivity on the comprehension of and attitudes toward online health content. *J Am Soc Inf Sci* 2007 Apr;58(6):766-776. [doi: [10.1002/asi.20557](https://doi.org/10.1002/asi.20557)]
35. Baumeister RF, Leary MR. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 1995 May;117(3):497-529. [Medline: [7777651](https://pubmed.ncbi.nlm.nih.gov/7777651/)]
36. Kelders SM. University of Twente. Enschede, The Netherlands: University of Twente; 2012. Understanding adherence to Web-based interventions [thesis] URL: <http://doc.utwente.nl/81967/> [accessed 2013-08-05] [WebCite Cache ID 6le8Ffe2A]
37. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement* 1977 Jun 01;1(3):385-401. [doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)]
38. Hayes SC, Strosahl K, Wilson KG. Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change. New York: Guilford Press; 1999.
39. Kabat-Zinn J. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. New York: Delacorte Press; 1990.
40. Kabat-Zinn J. Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life. New York: Hyperion; 1994.
41. Bohlmeijer ET, Hulsbergen M. Voluit Leven: Mindfulness of de kunst van het ervaren, nu als praktisch hulboek [Living to the Full: Mindfulness or the Art of Acceptance, Now As a Practical Help Book. Amsterdam, the Netherlands: Boom; 2008.

42. Bohlmeijer ET, Fledderus M, Rokx TA, Pieterse ME. Efficacy of an early intervention based on acceptance and commitment therapy for adults with depressive symptomatology: Evaluation in a randomized controlled trial. *Behav Res Ther* 2011 Jan;49(1):62-67. [doi: [10.1016/j.brat.2010.10.003](https://doi.org/10.1016/j.brat.2010.10.003)] [Medline: [21074752](https://pubmed.ncbi.nlm.nih.gov/21074752/)]
43. Fledderus M, Bohlmeijer ET, Pieterse ME, Schreurs KM. Acceptance and commitment therapy as guided self-help for psychological distress and positive mental health: a randomized controlled trial. *Psychol Med* 2012 Mar;42(3):485-495. [doi: [10.1017/S0033291711001206](https://doi.org/10.1017/S0033291711001206)] [Medline: [21740624](https://pubmed.ncbi.nlm.nih.gov/21740624/)]
44. Fledderus M, Bohlmeijer ET, Smit F, Westerhof GJ. Mental health promotion as a new goal in public mental health care: a randomized controlled trial of an intervention enhancing psychological flexibility. *Am J Public Health* 2010 Dec;100(12):2372. [doi: [10.2105/AJPH.2010.196196](https://doi.org/10.2105/AJPH.2010.196196)] [Medline: [20966360](https://pubmed.ncbi.nlm.nih.gov/20966360/)]
45. Kelders SM, Pots WT, Oskam MJ, Bohlmeijer ET, van Gemert-Pijnen JE. Development of a Web-based intervention for the indicated prevention of depression. *BMC Med Inform Decis Mak* 2013;13:26 [FREE Full text] [doi: [10.1186/1472-6947-13-26](https://doi.org/10.1186/1472-6947-13-26)] [Medline: [23425322](https://pubmed.ncbi.nlm.nih.gov/23425322/)]
46. Haringsma R, Engels GI, Beekman AT, Spinhoven P. The criterion validity of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of self-referred elders with depressive symptomatology. *Int J Geriatr Psychiatry* 2004 Jun;19(6):558-563. [doi: [10.1002/gps.1130](https://doi.org/10.1002/gps.1130)] [Medline: [15211536](https://pubmed.ncbi.nlm.nih.gov/15211536/)]
47. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med* 1997 Mar;27(2):363-370. [Medline: [9089829](https://pubmed.ncbi.nlm.nih.gov/9089829/)]
48. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983 Jun;67(6):361-370. [doi: [10.1111/j.1600-0447.1983.tb09716.x](https://doi.org/10.1111/j.1600-0447.1983.tb09716.x)]
49. Leary MR, Kelly KM, Cottrell CA, Schreindorfer LS, Mark R Leary, Duke University. 2007. Individual differences in the need to belong: mapping the nomological network [unpublished manuscript] URL: <http://people.duke.edu/~leary/scales.html> [accessed 2013-07-24] [WebCite Cache ID [6LbXv1mr](https://www.webcitation.org/6LbXv1mr)]
50. Nijland N, van Gemert-Pijnen JE, Boer H, Steehouder MF, Seydel ER. Increasing the use of e-consultation in primary care: results of an online survey among non-users of e-consultation. *Int J Med Inform* 2009 Oct;78(10):688-703. [doi: [10.1016/j.ijmedinf.2009.06.002](https://doi.org/10.1016/j.ijmedinf.2009.06.002)] [Medline: [19625210](https://pubmed.ncbi.nlm.nih.gov/19625210/)]
51. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. *J Med Internet Res* 2011;13(3):e71 [FREE Full text] [doi: [10.2196/jmir.1603](https://doi.org/10.2196/jmir.1603)] [Medline: [21959968](https://pubmed.ncbi.nlm.nih.gov/21959968/)]
52. Hirsh JB, Kang SK, Bodenhausen GV. Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. *Psychol Sci* 2012 Jun;23(6):578-581. [doi: [10.1177/0956797611436349](https://doi.org/10.1177/0956797611436349)] [Medline: [22547658](https://pubmed.ncbi.nlm.nih.gov/22547658/)]
53. Kaptein M, Eckles D. Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. In: *PERSUASIVE'10 Proceedings of the 5th international conference on Persuasive Technology*. Berlin/Heidelberg: Springer; 2010 Presented at: *PERSUASIVE'10: 5th international conference on Persuasive Technology*; June 7-10, 2010; Copenhagen, Denmark p. 82. [doi: [10.1007/978-3-642-13226-1_10](https://doi.org/10.1007/978-3-642-13226-1_10)]
54. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res* 2011;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
55. Fogg BJ. *Persuasive Technology: Using Computers to Change What We Think and Do*. Boston: Morgan Kaufmann; 2003.
56. Oinas-Kukkonen H, Harjumaa M. Communications of the Association for Information Systems. 2009. Persuasive systems design: key issues, process model, and system features URL: http://www.persuasive.eu/wp-content/uploads/2013/03/oinas_kukkonen_et_harjumaa_2009.pdf [accessed 2013-08-05] [WebCite Cache ID [6le95VZd2](https://www.webcitation.org/6le95VZd2)]
57. Crutzen R, Roosjen JL, Poelman J. Using Google Analytics as a process evaluation method for Internet-delivered interventions: an example on sexual health. *Health Promot Int* 2013 Mar;28(1):36-42. [doi: [10.1093/heapro/das008](https://doi.org/10.1093/heapro/das008)] [Medline: [22377974](https://pubmed.ncbi.nlm.nih.gov/22377974/)]
58. Tian H, Brimmer DJ, Lin JM, Tumpey AJ, Reeves WC. Web usage data as a means of evaluating public health messaging and outreach. *J Med Internet Res* 2009;11(4):e52 [FREE Full text] [doi: [10.2196/jmir.1278](https://doi.org/10.2196/jmir.1278)] [Medline: [20026451](https://pubmed.ncbi.nlm.nih.gov/20026451/)]
59. 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

- ACT:** Acceptance and Commitment Therapy
- ANOVA:** analysis of variance
- CES-D:** Center of Epidemiological Studies Depression scale
- HADS:** Hospital Anxiety and Depression Scale
- RCT:** randomized controlled trial
- SMS:** short message service

Edited by G Eysenbach; submitted 06.07.12; peer-reviewed by C Davies; comments to author 06.10.12; revised version received 18.02.13; accepted 13.06.13; published 20.08.13.

Please cite as:

Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JEW

Participants, Usage, and Use Patterns of a Web-Based Intervention for the Prevention of Depression Within a Randomized Controlled Trial

J Med Internet Res 2013;15(8):e172

URL: <http://www.jmir.org/2013/8/e172/>

doi: [10.2196/jmir.2258](https://doi.org/10.2196/jmir.2258)

PMID: [23963284](https://pubmed.ncbi.nlm.nih.gov/23963284/)

©Saskia M Kelders, Ernst T Bohlmeijer, Julia EWC Van Gemert-Pijnen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.08.2013. 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

Using Twitter to Examine Smoking Behavior and Perceptions of Emerging Tobacco Products

Mark Myslín¹; Shu-Hong Zhu², PhD; Wendy Chapman³, PhD; Mike Conway³, PhD

¹Department of Linguistics, University of California, San Diego, La Jolla, CA, United States

²Department of Family and Preventive Medicine, University of California, San Diego, La Jolla, CA, United States

³Department of Medicine, University of California, San Diego, La Jolla, CA, United States

Corresponding Author:

Mike Conway, PhD

Department of Medicine

University of California, San Diego

9500 Gilman Drive

La Jolla, CA, 92093

United States

Phone: 1 858 822 4931

Fax: 1 858 822 1934

Email: mconway@ucsd.edu

Abstract

Background: Social media platforms such as Twitter are rapidly becoming key resources for public health surveillance applications, yet little is known about Twitter users' levels of informedness and sentiment toward tobacco, especially with regard to the emerging tobacco control challenges posed by hookah and electronic cigarettes.

Objective: To develop a content and sentiment analysis of tobacco-related Twitter posts and build machine learning classifiers to detect tobacco-relevant posts and sentiment towards tobacco, with a particular focus on new and emerging products like hookah and electronic cigarettes.

Methods: We collected 7362 tobacco-related Twitter posts at 15-day intervals from December 2011 to July 2012. Each tweet was manually classified using a triaxial scheme, capturing genre, theme, and sentiment. Using the collected data, machine-learning classifiers were trained to detect tobacco-related vs irrelevant tweets as well as positive vs negative sentiment, using Naïve Bayes, k-nearest neighbors, and Support Vector Machine (SVM) algorithms. Finally, phi contingency coefficients were computed between each of the categories to discover emergent patterns.

Results: The most prevalent genres were first- and second-hand experience and opinion, and the most frequent themes were hookah, cessation, and pleasure. Sentiment toward tobacco was overall more positive (1939/4215, 46% of tweets) than negative (1349/4215, 32%) or neutral among tweets mentioning it, even excluding the 9% of tweets categorized as marketing. Three separate metrics converged to support an emergent distinction between, on one hand, hookah and electronic cigarettes corresponding to positive sentiment, and on the other hand, traditional tobacco products and more general references corresponding to negative sentiment. These metrics included correlations between categories in the annotation scheme ($\phi_{\text{hookah-positive}}=0.39$; $\phi_{\text{e-cigs-positive}}=0.19$); correlations between search keywords and sentiment ($\chi^2_4=414.50$, $P<.001$, Cramer's $V=0.36$), and the most discriminating unigram features for positive and negative sentiment ranked by log odds ratio in the machine learning component of the study. In the automated classification tasks, SVMs using a relatively small number of unigram features (500) achieved best performance in discriminating tobacco-related from unrelated tweets (F score=0.85).

Conclusions: Novel insights available through Twitter for tobacco surveillance are attested through the high prevalence of positive sentiment. This positive sentiment is correlated in complex ways with social image, personal experience, and recently popular products such as hookah and electronic cigarettes. Several apparent perceptual disconnects between these products and their health effects suggest opportunities for tobacco control education. Finally, machine classification of tobacco-related posts shows a promising edge over strictly keyword-based approaches, yielding an improved signal-to-noise ratio in Twitter data and paving the way for automated tobacco surveillance applications.

(*J Med Internet Res* 2013;15(8):e174) doi:[10.2196/jmir.2534](https://doi.org/10.2196/jmir.2534)

KEYWORDS

social media; twitter messaging; smoking; natural language processing

Introduction

Background

Social media platforms such as Twitter are rapidly becoming key resources for public health surveillance applications. Vast amounts of freely available, user-generated online content, in addition to allowing for efficient and potentially automated, real-time monitoring of public sentiment and informedness, allow for bottom-up discovery of emergent patterns that may not be readily detectable using traditional surveillance methodologies such as pre-formulated surveys. In this study, we demonstrate the feasibility of a Twitter-based “infoveillance” [1] methodology to monitor perceptions of tobacco usage, with a special focus on new public health challenges posed by hookah and electronic cigarettes (e-cigarettes). In particular, we collected a large corpus of tobacco-related Twitter posts, developed a specialized content analysis of these posts, and trained machine-learning algorithms to classify posts automatically according to their relevance to tobacco-related content categories.

Twitter and Public Health Surveillance

Twitter offers a number of key benefits as a data source for public health surveillance. First, the dataset is large and readily accessible. In 2012, 340 million tweets were being posted daily [2], and this content is freely available (albeit subject to legal restrictions on redistribution). Second, data may be automatically collected and analyzed in real time. Third, Twitter content is user-centric, thus reflecting trends that surveys may not capture or that users may not discuss in more formal contexts. Finally, Twitter demographics allow for greater representation of underserved and difficult-to-reach groups. African-American, Hispanic, younger, and urban populations are in fact overrepresented on Twitter relative to the general population [3]. Twitter use is most common among 18-29 (26% of whom use Twitter) and 30-49 year olds (14% of whom use Twitter). Men and women use Twitter in almost equal numbers. Twitter users reflect the general population in terms of education levels: 15% of US adults with Internet access use Twitter (8% on a typical day), and 9% of US adults use Twitter from their smartphones (and with the projected growth of smartphone use, this number is likely to increase) [4].

Recent applications have sought to harness the unique public health surveillance opportunities offered by Twitter. A number of studies have tracked public sentiment and informedness during natural disasters, such as the 2011 Tohoku earthquake [5] and various disease outbreaks [1,6-9]. High correlations are reported between Twitter statistics and Centers for Disease Control and Prevention (CDC) statistics with regard to influenza informedness, affirming the value of Twitter as a rapid, cost-effective health status surveillance methodology [10]. In a related vein, Twitter surveillance has revealed evidence of poor public informedness and misuse of antibiotics [11]. In addition to tracking public sentiment and information, Twitter has been used to monitor medical conditions, such as dental

pain in the community [12]. Finally, temporal patterns, such as those in problem drinking [13] and seasonal mood variation [14], have also been demonstrated using Twitter surveillance.

Twitter and Tobacco Surveillance

“Infoveillance” is defined by Eysenbach as “the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy” [1]. We believe that a Twitter-based infoveillance methodology can be profitably implemented in tobacco surveillance. Tobacco control is identified as a global public health priority by the World Health Organization [15], and tobacco use is the most preventable cause of disease in the United States. 400,000 smokers and former smokers die each year of smoking-related diseases, with an additional 38,000 nonsmokers dying prematurely due to second-hand smoke [16,17]. In this work, we are focused on using social media to analyze public perceptions of new and emerging tobacco control challenges, specifically hookah and e-cigarettes.

Preliminary tobacco research using Twitter data has addressed several specific domains. Freeman [18] analyzed use of social media by tobacco companies to promote smoking, while Lowe et al and Prochaska et al [19,20] studied Twitter-based smoking cessation networks. Finally, Prier et al [21] identified and classified tweets using automatic topic modeling. In our research, we built on this work in three key ways: (1) we developed a comprehensive, multidimensional content analysis, (2) we improved the signal-to-noise ratio in tobacco-related Twitter data by implementing machine learning classification techniques, and (3) we additionally focused on the special utility of Twitter surveillance for two new tobacco-related public health challenges, namely hookah and e-cigarettes. Note that the United States Federal Drug Administration (USFDA) has recently prioritized research on investigating public perceptions of new and emerging tobacco products, like hookah and e-cigarettes [22].

Objective

Our first objective was to provide a content analysis of tobacco-related tweets. Work reported by Prier et al [21] identifies major topic categories using Latent Dirichlet Allocation topic models [23]. The five emergent categories identified using the topic modeling methodology were tobacco abuse, tobacco cessation, promotion of bars and marijuana smoking, anti-smoking content, and a general, incohesive category. We built on this work by manually developing a comprehensive and multidimensional taxonomy of tobacco-related tweets that could then be employed in machine learning applications. Using the first four categories identified by Prier et al as a starting point, we used an iterative content analysis technique [24] to build a multidimensional annotation scheme of tobacco-related tweets reflecting those categories important for public health.

Our second objective was to improve the signal-to-noise ratio in Twitter data by automatically filtering out irrelevant content. Strictly keyword-based approaches are susceptible to lexical ambiguity in natural language: the keyword and wildcard combination *smok**, for example, matches not only tobacco-related tweets but also tweets referring to *smoked cheese*. In order to reduce the presence of this type of noise, we trained machine classifiers to distinguish between tobacco-related and unrelated tweets.

The third distinctive objective of our work was to demonstrate the utility of Twitter in addressing new public health challenges related to tobacco usage. Two such issues are the growing popularity of hookah and e-cigarettes. As we discuss below, Twitter surveillance is particularly suited to understanding these new challenges.

A hookah (also known as shisha or narghile) is a waterpipe used to smoke flavored tobacco. Hookah is smoked by an estimated 100 million people daily [25], with increasing numbers of users both in the United States and worldwide, especially among college-age adults [26-28]. While the health risks associated with hookah use are similar to those of cigarette smoking [29], perceptions are widespread that hookah is safer [28,30]. Further, waterpipe usage is subject to fewer regulations than cigarettes, with frequent exemptions on bans in bars [31]. Furthermore, hookah products are easily accessible via Internet marketing sites and in venues that do not verify age [28]. Despite the growing list of health concerns associated with hookah, no interventions have been designed specifically for this form of tobacco use [32]. Its growing popularity among young users and its widespread availability online make Twitter a key resource for its surveillance.

E-cigarettes (or e-cigs) are another recently popular tobacco product subject to only sparse regulation and research. An e-cigarette is an electronic inhaler that produces vapor to simulate cigarette smoking and that may or may not contain nicotine. While e-cigarettes have surged in popularity as cessation devices, no consensus exists among public health researchers regarding their health effects, and they are not endorsed by either the USFDA or the CDC [33]. Indeed, some researchers show that e-cigarettes carry health risks and could appeal to nonsmokers, especially children, due to their novelty, flavorings, and possibly overstated claims of safety [34].

Regulation of e-cigarettes is sparse and variable by jurisdiction—no warning labels are required, and the product is easily available online [35]. Indeed, online marketing has surged [20], and by September 2010, Google searches related to e-cigarettes were several-hundred-fold more frequent than those related to cessation medications [36]. The centrality of the Internet to the rise of e-cigarettes underscores the value of Twitter surveillance of this product.

Methods

Data Collection

Using the Twitter Application Programming Interface (API), we collected a sample of tweets between November 2011 and July 2012 that represented 1% of the entire Twitter feed. This

1% sample consisted of an average 1.3 million tweets per day. In order to extract tobacco-related tweets from this dataset, we constructed a list of keywords relevant to general tobacco usage as well as hookah and e-cigarettes. Our initial list consisted of 30 such terms culled from online slang dictionaries, but we pruned this list to the 11 terms that were attested more than once per day in our Twitter sample (see below). These were *cig**, *nicotine*, *smok**, *tobacco*; *hookah*, *shisha*, *waterpipe*; *e-juice*, *e-liquid*, *vape*, and *vaping* (where * is a wildcard such that *cig** matches tweets containing *cigar*, *e-cig*, and so on).

Our initial dataset included all tweets containing these keywords at 15-day intervals from December 5, 2011, to July 17, 2012, inclusive, which results in equal sampling of each day of the week. We thus avoided potential bias based on day of the week, which has been observed for alcohol-related tweets, which spike in positive sentiment on Fridays and Saturdays [13]. For each of the 16 days resulting from our sampling technique, all tweets matching any of the listed keywords were included. Tweets matching these tobacco-related keywords reflected 0.17% of all tweets in the Twitter API 1% sample. The vast majority of these keyword-relevant tweets corresponded to unique Twitter users: on average, each username was associated with 1.07 tweets in the sample.

One of our keywords, *smok**, was dramatically more frequent and ambiguous than any of the others, matching far more tobacco-irrelevant tweets (for example, tweets referring to *smoked cheese*). In a preliminary sample of 500 *smok** tweets, only 16.8% were relevant to tobacco according to manual classification. Furthermore, over 100,000 *smok** tweets were included in our 16-day dataset, making hand classification impractical. We thus included *smok** tweets only for days where there were less than 400 total tweets matched by all other keywords, so that each day's total tweet count was at least 400, ensuring a balance such that no individual date was underrepresented. Following this procedure, 0.04% of all *smok** tweets were included in the dataset. The resulting final dataset thus contained 7362 tweets, with a mean of 460 tweets per day (SD 35).

Manual Content and Sentiment Analysis

We developed a triaxial classification scheme to capture each tweet's genre, theme, and sentiment. The former two axes are similar in scope to the content and qualifier categories developed in Chew & Eysenbach [7]. Genre reflects the format of the tweet (for example, formulaic joke, news item, or personal experience), and theme reflects the domain of the actual content conveyed (including such categories as health issues, underage usage, and tobacco policy). Sentiment, the third axis of the coding scheme, simply encodes the stance expressed in the tweet toward tobacco or its users, whether positive, negative, or neutral. Categories within each of the three axes were developed iteratively on the basis of a separate pilot dataset of approximately 1000 tweets from another date, which 2 annotators (authors MM and MC) classified according to an early version of the scheme. Upon review and discussion, several overly broad categories were split, while sparse, related categories were collapsed. A final version of the coding scheme was adopted when interannotator agreement among the 2

annotators on a set of 150 tweets exceeded a kappa level of 0.7 for each of the three axes. A complete list of all categories within this scheme is available in [Figure 1](#), and detailed descriptions and example tweets for each category are presented in [Multimedia Appendix 1](#). [Figure 2](#) shows two examples of how tweets are classified using the annotation scheme.

The set of 7362 tweets was then manually classified according to the final version of this scheme by the 2 annotators. Tweets were assigned multiple categories within a single axis if applicable, and duplicate or re-tweeted posts were included only once to prevent spam or overly popular posts from biasing the sample. Non-English, unintelligible, or tobacco-irrelevant tweets were coded as belonging to none of the categories in the classification scheme.

Intercategory Correlations

In order to discover emergent trends in tobacco-related Twitter content, we computed correlations for each pairwise combination of the 30 categories within the entire coding scheme. In other words, given two categories such as *hookah* and *positive sentiment*, we compared the number of tweets manually classified under both categories to the number expected by chance to be classified under both categories. The contingency coefficient phi (which is equivalent to Cramer's V in the current 2×2 case) equals the square root of χ^2/n , where χ^2 is the chi-square statistic for the 2×2 contingency table, and n is the total number of observations. The phi coefficient ranges from 0 to 1, with 0 indicating no correlation between the two categories and 1 indicating perfect correlation.

Machine Learning

We compared the performance of several machine learning algorithms on three classification tasks on the corpus of manually annotated tweets: relevance to tobacco, positive sentiment, and negative sentiment. Relevance to tobacco was operationalized as whether the tweet was classified under any of the categories in the scheme. Our goal was to test the feasibility of creating a natural language processing machine learning classifier with which we could automatically identify tobacco-related tweets in real-time.

We varied three parameters for each task: the machine learning algorithm, the order of n-gram used as features, and the number of features used. Algorithms used were Naïve Bayes, k-nearest-neighbors (KNN), and Support Vector Machines (SVM) (see [Figure 3](#) for a brief description of these algorithms)

[37], and features were either unigrams, bigrams, or trigrams (see [Figure 4](#) for a description of n-grams). The number of features ranged from 1 to the number of unique n-grams present for the current task, tested at approximately logarithmic intervals from 1 to 1000 and at intervals of 500 thereafter. Feature selection was determined by information gain, which measures the increase in bits of information when a term is present versus absent (see, for example, Yang & Pedersen for discussion [38]). Comparative studies of feature selection metrics report information gain as one of the best-performing metrics for text classification [38,39]. Our goal in using feature selection was twofold. First, we wanted to identify those words and phrases most associated with tobacco-related tweets. Second, we wanted to use these high-quality features in order to increase the classification accuracy of our machine learning classifiers.

We employed the Rainbow toolkit [40] to train and test the above classifiers and manually implemented a 10-fold cross-validation routine for each classification task using the hand-annotated dataset. In cross-validation, the entire hand-annotated dataset is broken into k equally sized folds, and the classification task is performed k times, each time with a different fold held out as test data, and all other folds included as training data. Classification accuracies for each of the k iterations are then simply averaged.

Features used for machine learning were represented as binary presence/absence of words in a tweet rather than the number of times each term occurred in a tweet. Term frequencies are unlikely to be significantly more informative, since words are relatively rarely repeated within tweets (mean type-token ratio 0.96, SD 0.08). Two additional standard feature-processing measures were taken: first, all tweets were passed through the Porter stemmer [41], which converts words (such as *smoked* and *smoking*) to their bare stems (in both cases here, *smok*), so that different conjugations of the same lexical item are not counted as distinct features. Second, extremely frequent function words such as *the* and *is*, which are unlikely to be relevant to the classification task, were excluded as features using the standard 524-word SMART stoplist [42]. All other machine learning parameters were Rainbow defaults. [Figure 5](#) summarizes the machine learning workflow.

In order to evaluate the machine learning results, five standard classification metrics were computed for each task. Accuracy is simply the percentage of tweets correctly classified by the algorithm. We also computed precision, recall, specificity, and F scores, which are defined in [Multimedia Appendix 2](#).

Figure 2. Example tweets manually classified using annotation scheme (relevant categories are shaded).

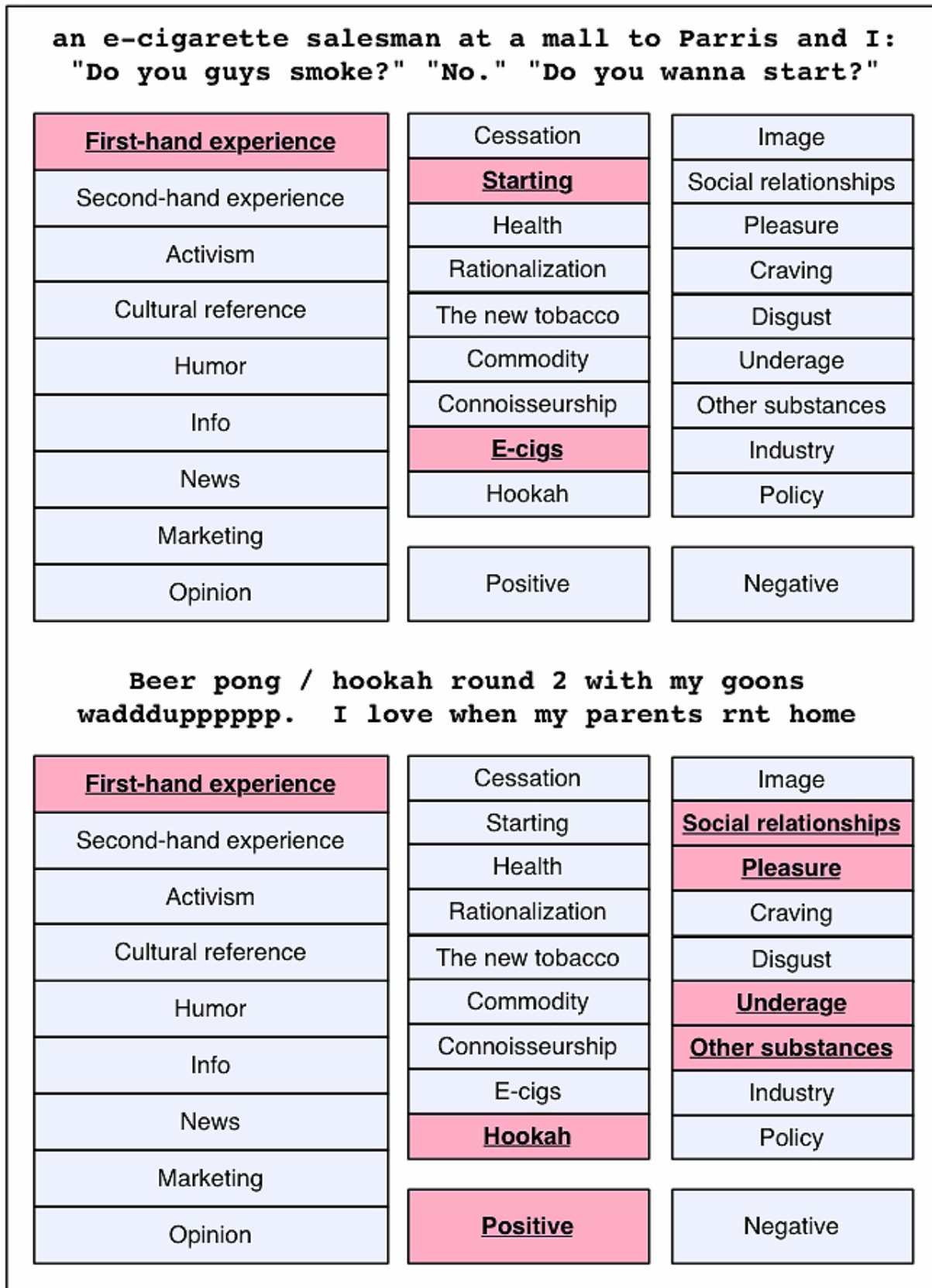


Figure 3. Machine learning algorithm description.

Naive Bayes Algorithm
The Naive Bayes algorithm is popular in text classification due to its relative simplicity and computational efficiency. The algorithm is <i>bayesian</i> in that it relies on Bayes' method for calculating probability, and <i>naive</i> in that it assumes that all features (in our case, ngrams) are conditionally independent (and equally important) with respect to the classification task.
Support Vector Machine Algorithm
The Support Vector Machine algorithm is a sophisticated classification algorithm developed in the 1990s that has become popular in text classification due to its high accuracy for many classification tasks. Support Vector Machines achieve such high performance by determining a hyperplane that is optimally discriminating with respect to the two classes of interest.
K-Nearest-Neighbour Algorithm
Like the Naive Bayes algorithm, the K-Nearest-Neighbour algorithm is conceptually relatively simple, but unlike Naive Bayes, it is relatively computationally inefficient. For each document (or in our case, tweet) in the dataset, the algorithm identifies the <i>k</i> nearest (i.e. most similar) examples in the labeled training set, determines the majority class among these <i>k</i> examples, and then simply assigns the document the majority class.

Figure 4. N-gram text representation.

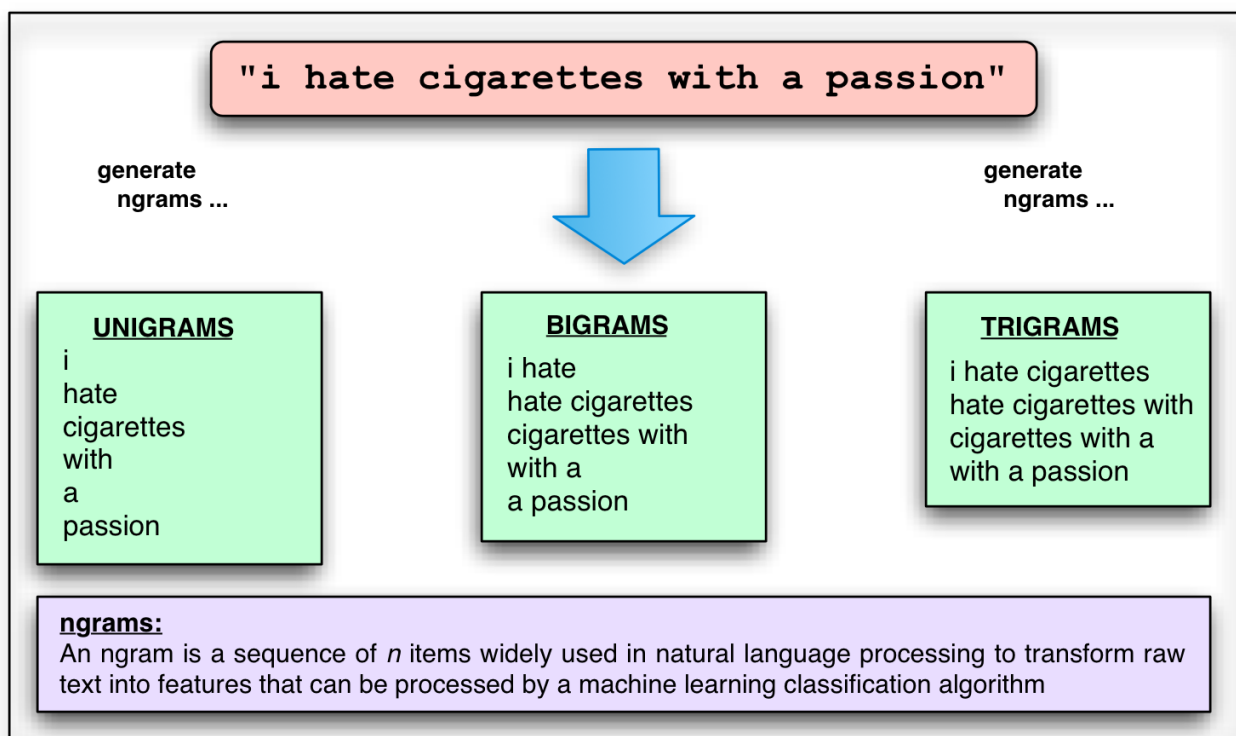
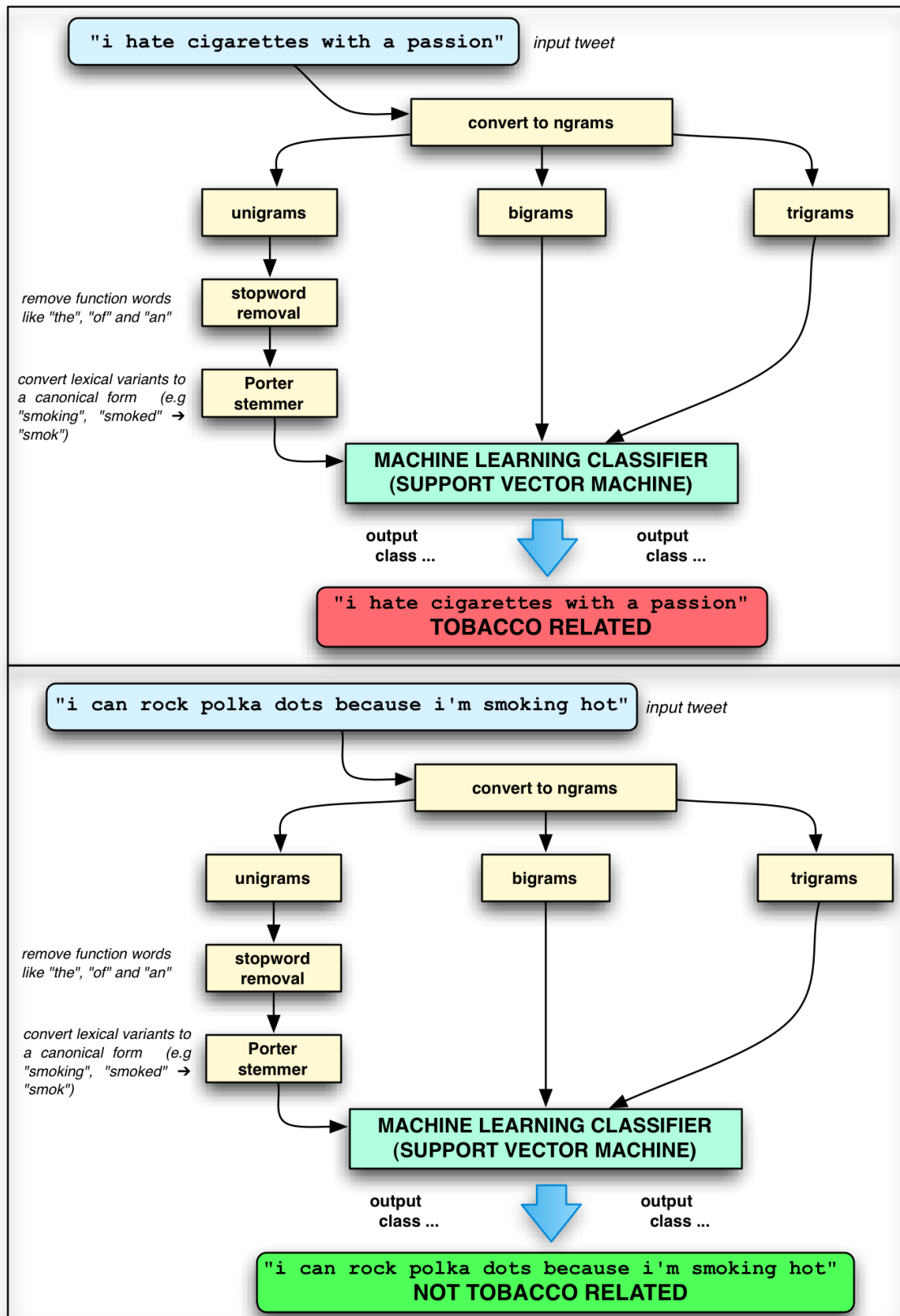


Figure 5. Machine learning experiment workflow.



Results

Content and Sentiment Analysis

The corpus of 7362 tweets was annotated by authors MM and MC according to the classification scheme described in the Methods section. Interannotator agreement (κ) met the standard threshold of 0.7 for each of the three axes of the scheme: genre=0.78, theme=0.70, sentiment=0.77. Of the tweets, 4215 (57.3%) were classified as relevant to tobacco, with the remainder comprising tweets that were not in English or that matched alternate senses of their keyword, such as *smoked cheese* in the case of *smok**.

Among the tobacco-related tweets (ie, 4215 out of a total of 7362), the most prevalent genre was first-hand experience, matching 40% of tweets, followed by second-hand experience (14%), and opinion (9%) (recall that tweets may be assigned multiple categories). The top themes were hookah (20%), cessation (14%), and pleasure (11%). Finally, sentiment toward tobacco was overall more positive (46% of tweets) than negative (32%) or neutral, even excluding the 9% of tweets categorized as marketing, which resulted in a 41%/30% positive/negative ratio.

Search keywords associated with each tweet correlated significantly with more general properties, such as sentiment. Examining the five most frequent keywords (representing 96% of tweets), [Figure 6](#) illustrates the tendency for tweets matching the keywords *hookah*, *shisha*, and *vape/vaping* to be classified as showing positive sentiment more often than expected by chance, and for those matching *tobacco* to show negative sentiment disproportionately often (note that low frequency keywords—*nicotine*, *waterpipe*, *e-juice*, and *e-liquid*—were excluded). The correlation is highly significant according to a two-tailed chi-square test for independence ($\chi^2_4 = 414.50$, $P < .001$, Cramer's $V = 0.36$). In this way, a general split in sentiment is observed between, on one hand, the new public health challenges represented by hookah and e-cigarettes, which are viewed more positively, and on the other hand, traditional products such as cigarettes as well as more general references to tobacco, which are viewed more negatively. In other words, smoking hookah is viewed more favorably than smoking traditional tobacco products, even though smoking hookah typically involves smoking tobacco.

Intercategory Correlations

Correlations between all pairwise combinations of categories in the classification scheme, computed as described in the Methods section, are reported in [Figure 1](#). The highest intercategory correlations were observed between (1) underage usage and social image (0.6), (2) e-cigarettes and marketing (0.54), and (3) positive sentiment and first-person experience (0.47).

Machine Learning

The three classification tasks investigated here are (1) relevance to tobacco, (2) positive sentiment toward tobacco, and (3) negative sentiment toward tobacco. For each task, we varied n-gram size, number of features, and machine learning

algorithm. In all tasks, unigram feature sets yielded consistently better performance than bigrams on all measures except recall, and bigrams similarly generally outperformed trigrams. A relatively small feature set was generally optimal: [Figure 7](#) illustrates that for the tobacco-relevance task, classification accuracy peaks or levels off with well under 5000 features, and indeed maximum classification accuracy is achieved by a classifier trained on 500 features. SVMs generally yielded the best performance, followed by Naïve Bayes and KNN algorithms, respectively. In discriminating between tobacco-related and -unrelated tweets in order to improve the signal-to-noise ratio in Twitter data, a substantial improvement (82% classification accuracy) over the majority-class baseline (57% classification accuracy) was achieved. [Table 1](#) summarizes performance results for each task using 500 features.

The most informative unigram features for each of the three classification tasks, ranked by log odds ratio, are listed in [Table 2](#). Among the most informative features distinguishing tobacco-related from unrelated tweets are relatively predictable, unambiguous words such as *cigarette*, *hookah*, and *tobacco*. Several other emergent classes of words are apparent: marketing-related words including *buy* and *http* (typically part of sales website URLs); words semantically or pragmatically associated with tobacco usage such as *smell* and *bar*; and conversational words such as *I'm*, *don't*, and *lol* that are suggestive of personal expression rather than, for example, news or marketing.

Turning to the most informative features for positive and negative sentiment, several Twitter-specific expressions appear. *gt* and *lt* correspond to the greater-than symbol and the less-than symbol, which are, respectively, explicit tokens of positive and negative sentiment. *smh*, an acronym for *shaking my head*, is a general token of disapproval and is among the most informative features for negative sentiment toward tobacco.

A key point of contrast between highly informative positive words and highly informative negative words is evident in the kind of tobacco product to which they refer. Words related to hookah and e-cigarettes are highly predictive of positive sentiment (respectively, *hookah*, *hose*, *shisha*; *electronic*), whereas *cigarettes* and more general terms such as *smoke* and *tobacco* are predictive of negative sentiment. Discussion of this distinction, as well as its relation to the similar result in the interaction of search keywords and sentiment, is taken up in the next section.

The remaining positive and negative unigrams reveal informative semantic groupings. Words related to recreation and social interaction generally predict positive sentiment toward tobacco, and include *bar*, *tonight*, and *night*. Marketing-related words, such as *buy*, *free*, *coupon*, *checkout*, *code*, and *win*, are also prevalent in the positive category. Groupings in the negative category include words related to disgust and social image, such as *nasty*, *unattractive*, *people*, and *girls*, where these last two terms most often occurred in tweets disapproving of particular social groups' use of tobacco. Finally, words predictive of negative sentiment toward tobacco were also related to health, information, and cessation: *health*, *kill*, *study*, *finds*, *quit*.

Figure 6. Tweet sentiment by search keyword.

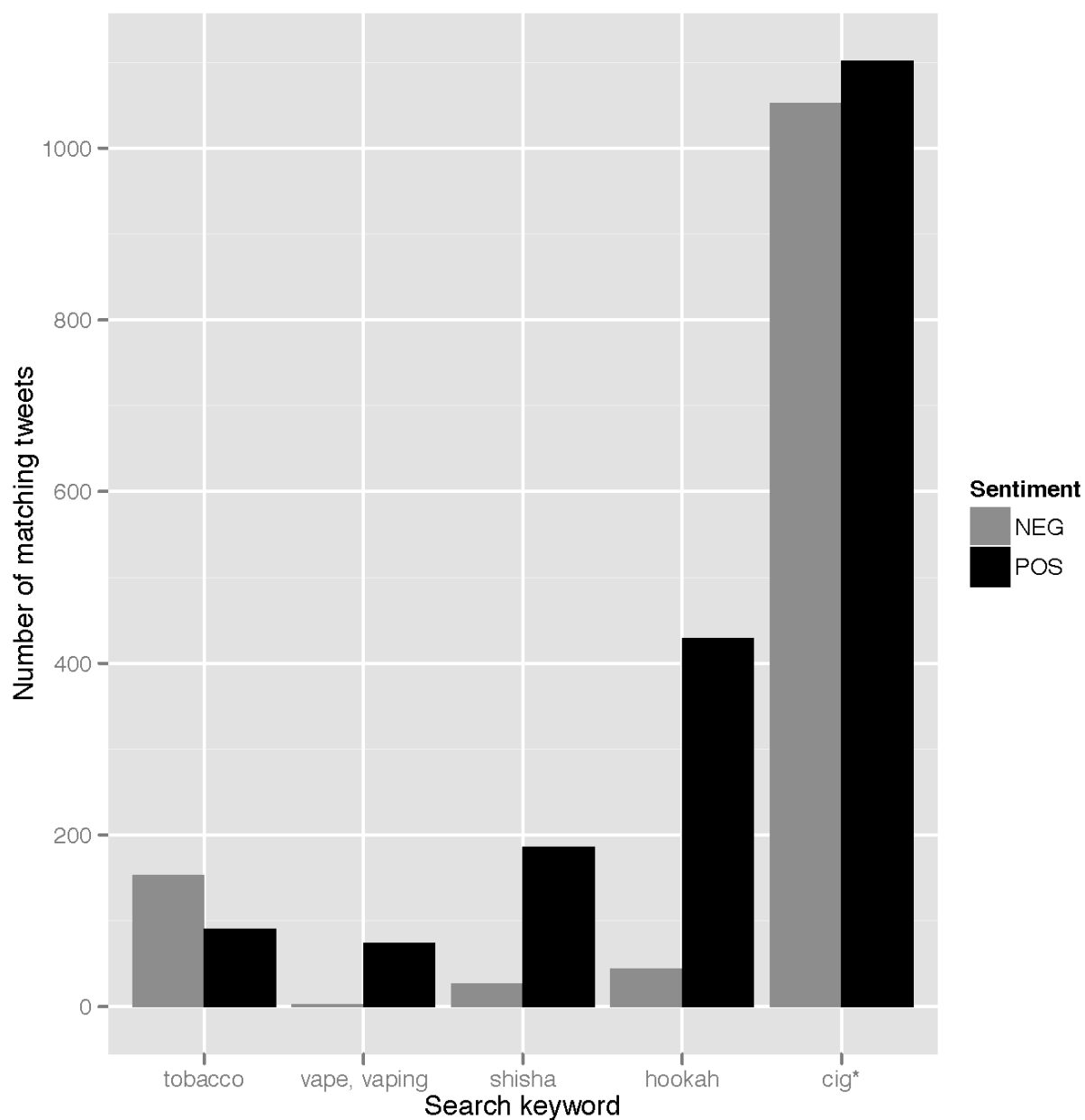


Table 1. Performance measures for tobacco relevance, positive sentiment, and negative sentiment classification tasks using 500 features (baseline classification accuracies [majority class] are 57% for relevance, 74% for positive sentiment, and 82% for negative sentiment).

Features	Naïve Bayes					KNN					SVM				
	Acc ^a	F	Pre ^b	Rec ^c	Spe ^d	Acc	F	Pre	Rec	Spe	Acc	F	Pre	Rec	Spe
Relevance															
Unigrams	0.77	0.83	0.73	0.95	0.53	0.73	0.78	0.73	0.83	0.59	0.82	0.85	0.82	0.88	0.75
Bigrams	0.66	0.77	0.63	0.97	0.24	0.65	0.76	0.63	0.97	0.24	0.73	0.75	0.82	0.69	0.79
Trigrams	0.61	0.74	0.6	0.99	0.1	0.6	0.74	0.59	0.97	0.11	0.61	0.74	0.59	0.99	0.1
Positive sentiment															
Unigrams	0.76	0.5	0.56	0.45	0.87	0.76	0.37	0.58	0.27	0.93	0.75	0.38	0.53	0.3	0.91
Bigrams	0.77	0.44	0.62	0.34	0.93	0.76	0.42	0.58	0.33	0.92	0.77	0.43	0.61	0.33	0.92
Trigrams	0.76	0.26	0.62	0.16	0.96	0.76	0.26	0.62	0.17	0.96	0.76	0.27	0.61	0.17	0.96
Negative sentiment															
Unigrams	0.84	0.52	0.57	0.48	0.92	0.72	0.3	0.27	0.33	0.8	0.83	0.39	0.53	0.3	0.94
Bigrams	0.85	0.35	0.73	0.23	0.98	0.31	0.3	0.18	0.82	0.2	0.84	0.44	0.59	0.35	0.95
Trigrams	0.84	0.24	0.76	0.14	0.99	0.22	0.3	0.18	0.94	0.07	0.84	0.37	0.66	0.25	0.97

^aAcc: accuracy.

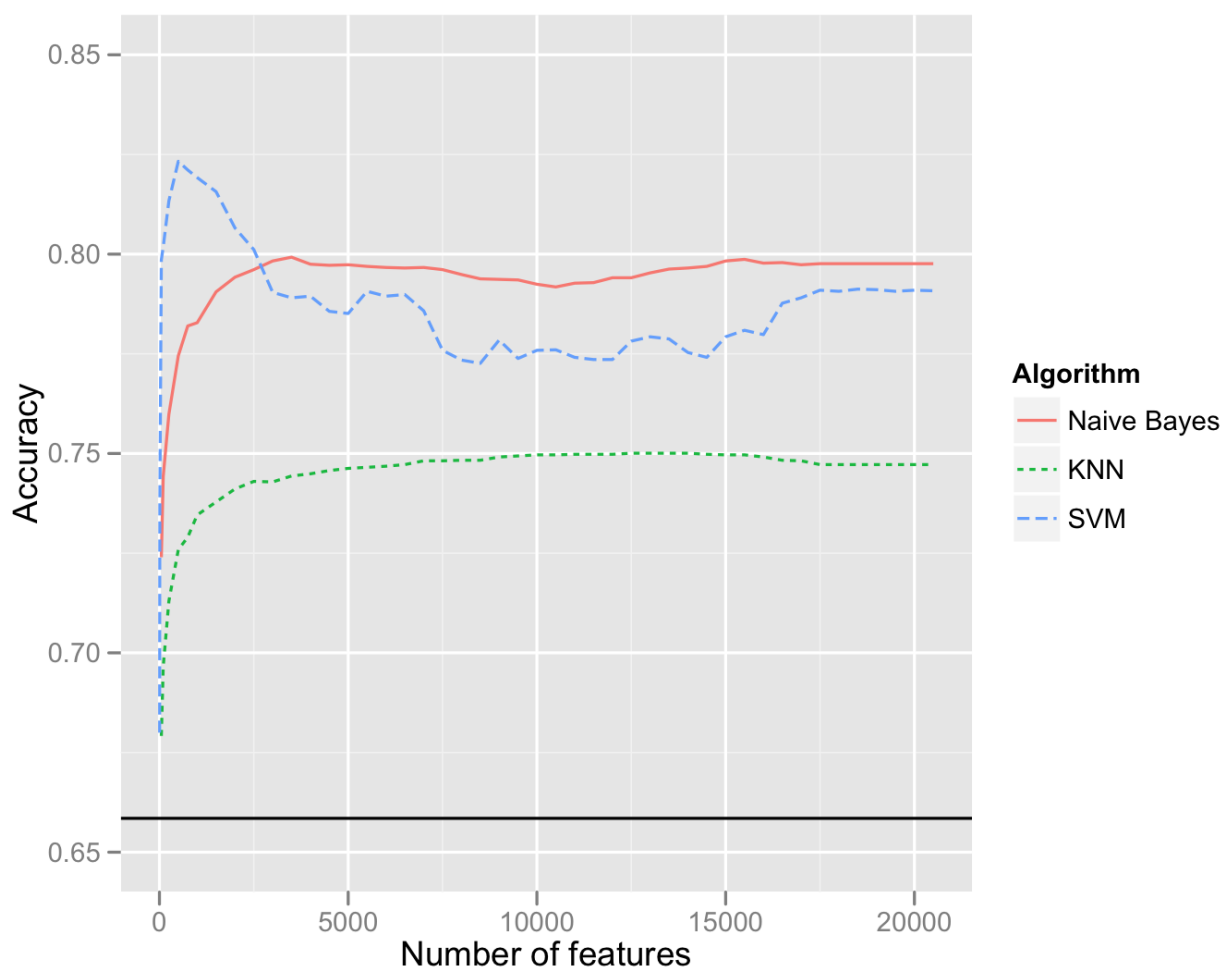
^bPre: precision.

^cRec: recall.

^dSpe: specificity.

Table 2. Most discriminating unigram features for tobacco-related, positive sentiment, and negative sentiment categories, ranked by log odds ratio.

Tobacco-related	Positive sentiment	Negative sentiment
cigarette	hookah	It
hookah	cigar	cigarettes
It	bar	smell
smoking	tonight	hate
tobacco	gt	smoke
cigs	electronic	people
electronic	night	disgusting
http	good	tobacco
smell	code	finds
cigar	checkout	study
im	love	girls
bar	lol	alcohol
hate	free	nasty
day	ecigarette	unattractive
dont	buy	smh
gt	hose	smells
buy	win	kill
lol	coupon	health
people	flavored	mouth
good	shisha	quit

Figure 7. Classification accuracy as a function of number of unigram features for 3 algorithms in the tobacco-relevance task.

Discussion

Principal Findings

The Twitter surveillance results converge in several key classes of findings, which we discuss in turn in this section. First, the content analysis allows for a general pulse or snapshot to be taken of tobacco-related discussion on Twitter. Second, new insight can be gained into causes for positive and negative sentiment toward tobacco, especially with respect to hookah and e-cigarettes. Finally, several potential opportunities for tobacco education emerge, and we discuss them in the context of future research directions.

The relative prevalence of the various categories in the content analysis reflect a general pulse of tobacco-related discussion on Twitter. By far the most common categories are personal experiences and opinion, affirming the value of Twitter in assessing public sentiment and informedness. The next most common genre, marketing, is followed relatively distantly by information and news, and most tweets in these categories are not posted by recognized health or news organizations. In sum, reliable information is far less accessible on Twitter than are opinions, marketing posts, and information from unverified sources, indicating potential for greater public education in tobacco prevention policies.

Twitter surveillance allows for new insight into the correlates of positive and negative sentiment toward tobacco. Among Twitter users that post about tobacco in our dataset, sentiment is overall more positive than negative, even with marketing posts excluded. The strongest correlate of positive sentiment is first-hand personal experience, while negative sentiment correlates more strongly with opinion. In this regard, Twitter surveillance may reveal insights not available through surveys, where participants do not spontaneously relate experiences to an audience of friends and followers and are instead more likely to express more carefully crafted opinions. Indeed, surveys may thus underestimate the prevalence of positive sentiment toward tobacco.

Among the clearest correlates of positive sentiment are hookah and e-cigarettes. On all measures computed in this study, including (1) correlations between categories in the annotation scheme, (2) correlations between search keywords and sentiment, and (3) most discriminating unigram features for positive and negative sentiment, a split emerged between, on one hand, hookah and e-cigarettes as corresponding to positive sentiment, and on the other hand, other products as well as general references to tobacco corresponding to negative sentiment. Especially in the case of hookah, such a split may indicate a disconnect in public perception between popular tobacco products and risk factors associated with tobacco use.

in general, presenting a distinct opportunity for outreach and education by tobacco control organizations.

Social relationships, especially among younger users, emerge as another key component of positive sentiment toward tobacco on Twitter, often in conjunction with products such as hookah. In the following example, tobacco usage is a central component of a positive experience in a social relationship: “Smoking that good hookah with the bro Sultan! #GoodOldDays #brotherforlife”. These positive tobacco-centric social experiences also frequently involve young or under-age users: “Beer ponggg / hookah round 2 with my goons waddduppppppp. I love when my parents rnt home!”

In a related vein, these products are also associated with initiation of tobacco usage, as in the following: “an e-cigarette salesman at a mall to Parris and I: ‘Do you guys smoke?’ ‘No.’ ‘Do you wanna start?’.”

In this way, positive sentiment toward tobacco appears to participate in a complex interaction between newer products such as hookah and e-cigarettes, younger users, and positive social experiences.

A social component is also central to negative sentiment toward tobacco. Categories corresponding to disgust and stereotypes were among the most highly correlated with negative sentiment, in fact outranking the explicit health category. A key distinction, however, is that while the category of social image correlated with negative sentiment, social relationships correlated with positive sentiment. Taken together, these findings indicate that social factors are central in driving sentiment toward tobacco and suggest that public health campaigns may do well to make use of this correlation.

Several novel findings, in sum, speak to the unique insights available through Twitter surveillance. All measures converged on an emergent distinction between two recently popular tobacco products, hookah and e-cigarettes, which corresponded to positive sentiment, and other products as well as references to tobacco more generally, which corresponded to negative sentiment. Sentiment toward tobacco overall among Twitter users is more positive than negative, affirming Twitter’s value as a resource to understand positive sentiment in developing

improved prevention policies. Negative sentiment is equally useful; for example, observed high correlations between negative sentiment and social image, but not health issues, may usefully inform tobacco control strategies. Twitter surveillance further reveals opportunities for education. Positive sentiment toward the term *hookah* but negative sentiment toward *tobacco* suggests a disconnect in users’ perceptions of the health effects of hookah (ie, hookah is not regarded in the same negative light as traditional tobacco products). Finally, machine classification of tobacco-related posts shows a promising edge over strictly keyword-based approaches, yielding an improved signal-to-noise ratio and paving the way for automated tobacco surveillance applications.

Limitations

The work reported in this paper does have some limitations. First, we harvested our data from the free 1% Twitter feed, rather than the full Twitter firehose. Second, our annotated dataset was relatively small, and there is some risk of our model overfitting. Third, the number of smoking keywords used to identify tobacco-relevant tweets was quite limited. It would be useful to augment our keyword list with tobacco-related slang (eg, “cancer sticks”, “coffin nails”) or electronic cigarette brands (eg, “blucigs”, “greensmoke”). Fourth, in this work we have concentrated exclusively on analyzing tobacco-related tweets using natural language processing rather than on the social network aspect of Twitter (ie, we did not analyze the characteristics of those tweets most likely to be retweeted). Finally, one key issue that we have not addressed in this work is the role of novelty effects in attitudes towards e-cigarettes (ie, will interest in the products be sustained over time?). In future work we will address all these issues.

Our medium-term goal, building on the work described in this paper, is to create a Web-based social media monitoring system for tobacco-related products and smoking behaviors, integrating natural language processing, geographical information systems, and social network analysis to provide a service that will allow public health workers and other interested parties to monitor and track public attitudes towards a range of both established and emerging tobacco products, and to formulate policy and interventions accordingly.

Acknowledgments

We would like to thank Ms Madeleine Lee (Department of Family & Preventive Medicine at the University of California, San Diego) for her support and useful suggestions and Drs Son Doan and Sharon Cummins (Department of Medicine at the University of California, San Diego, and the Department of Family & Preventive Medicine at University of California, San Diego, respectively) for offering useful comments on an earlier version of this manuscript.

This work was supported in part by grants from the National Cancer Institute (grant: U01 CA154280) and the NIH Roadmap for Medical Research (grant: U54HL108460). Author MM was partially supported by a Jacob K Javits Graduate Fellowship.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Annotation scheme.

[PDF File (Adobe PDF File), 144KB - [jmir_v15i8e174_app1.pdf](#)]

Multimedia Appendix 2

Evaluation metrics.

[PDF File (Adobe PDF File), 61KB - [jmir_v15i8e174_app2.pdf](#)]

References

1. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](#)] [Medline: [19329408](#)]
2. Twitter turns six. URL: <http://blog.twitter.com/2012/03/twitter-turns-six.html> [accessed 2012-10-13] [WebCite Cache ID [6BOdO7FWf](#)]
3. Mislove A, Lehmann S, Ahn Y, Onnela J, Rosenquist J. Understanding the demographics of Twitter users. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media. 2011 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona.
4. Smith A, Brenner J. Twitter use 2012.: Pew Internet Research URL: <http://pewinternet.org/Reports/2012/Twitter-Use-2012.aspx> [accessed 2012-10-13] [WebCite Cache ID [6BOdGaoJh](#)]
5. Doan S, Vo B, Collier N. An analysis of Twitter messages in the 2011 Tohoku Earthquake. 2011. URL: <http://www.informatik.uni-trier.de/~ley/db/conf/ehealth/ehealth2011.html> [accessed 2012-10-14] [WebCite Cache ID [6BP56TaFs](#)]
6. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. *J Biomed Semantics* 2011;2 Suppl 5:S9 [FREE Full text] [doi: [10.1186/2041-1480-2-S5-S9](#)] [Medline: [22166368](#)]
7. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS One* 2010;5(11):e14118 [FREE Full text] [doi: [10.1371/journal.pone.0014118](#)] [Medline: [21124761](#)]
8. Kriek M, Dreesman J, Otrusina L, Denecke K. A new age of public health: identifying disease outbreaks by analyzing tweets. In: Proceedings of Health WebScience Workshop, ACM Web Science Conference. 2011 Presented at: ACM Web Science Conference; June 14-17, 2011; Koblenz, Germany.
9. Paul M, Dredze M. You are what you tweet: analyzing Twitter for public health. In: Proceedings of the 5th International AAAI Conference on Weblogs and Social Media. 2011 Presented at: 5th International AAAI Conference on Weblogs and Social Media; July 17-21, 2011; Barcelona.
10. Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. In: Proceedings of the First Workshop on Social Media Analytics. 2010 Presented at: First Workshop on Social Media Analytics; July 25-28, 2010; Washington, DC.
11. Scanfeld D, Scanfeld V, Larson EL. Dissemination of health information through social networks: twitter and antibiotics. *Am J Infect Control* 2010 Apr;38(3):182-188 [FREE Full text] [doi: [10.1016/j.ajic.2009.11.004](#)] [Medline: [20347636](#)]
12. Heavilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. *J Dent Res* 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: [10.1177/0022034511415273](#)] [Medline: [21768306](#)]
13. West J. Temporal variability of problem drinking on Twitter. *Open Journal of Preventive Medicine* 2012;2(1):43-48 [FREE Full text]
14. Golder SA, Macy MW. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 2011 Sep 30;333(6051):1878-1881 [FREE Full text] [doi: [10.1126/science.1202775](#)] [Medline: [21960633](#)]
15. World Health Organization. WHO Report on the global tobacco epidemic 2011: warning about the dangers of tobacco. 2011. URL: http://www.who.int/tobacco/global_report/2011/en/ [accessed 2012-10-14] [WebCite Cache ID [6BP5YVMi4](#)]
16. Armour B, Woolery T, Malarcher A, Pechacek T, Husten C. Annual smoking- attributable mortality, years of potential life lost, and productivity losses. *Morbidity and Mortality Weekly Report* 2005;54(25):625-628 [FREE Full text]
17. US Department of Health and Human Services. The health consequences of smoking: a report for the Surgeon General. 2004. URL: <http://www.surgeongeneral.gov/library/reports/smokingconsequences/index.html> [accessed 2012-10-14] [WebCite Cache ID [6BP5nq84N](#)]
18. Freeman B. New media and tobacco control. *Tob Control* 2012 Mar;21(2):139-144. [doi: [10.1136/tobaccocontrol-2011-050193](#)] [Medline: [22345236](#)]
19. Lowe JB, Barnes M, Teo C, Sutherns S. Investigating the use of social media to help women from going back to smoking post-partum. *Aust N Z J Public Health* 2012 Feb;36(1):30-32. [doi: [10.1111/j.1753-6405.2012.00826.x](#)] [Medline: [22313703](#)]
20. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. *Tob Control* 2012 Jul;21(4):447-449 [FREE Full text] [doi: [10.1136/tc.2010.042507](#)] [Medline: [21730101](#)]
21. Prier K, Smith M, Giraud-Carrier C, Hanson C. Identifying health related topics on Twitter. *Social Computing, Behavioral-Cultural Modeling and Prediction* 2011:18-25.
22. U.S. Food and Drug Administration RFA. URL: <http://grants.nih.gov/grants/guide/rfa-files/RFA-DA-13-003.html> [accessed 2012-12-20] [WebCite Cache ID [6D40gp4Ss](#)]
23. Blei D, Ng A, Jordan M. Latent Dirichlet allocation. *J Mach Learn Res* 2003;3:993-1022.

24. Neuendorf K. The Content Analysis Guidebook. London: SAGE Publications; 2001.
25. Rogers JM. Tobacco and pregnancy: overview of exposures and effects. Birth Defects Res C Embryo Today 2008 Mar;84(1):1-15. [doi: [10.1002/bdrc.20119](https://doi.org/10.1002/bdrc.20119)] [Medline: [18383133](https://pubmed.ncbi.nlm.nih.gov/18383133/)]
26. Cobb C, Ward KD, Maziak W, Shihadeh AL, Eissenberg T. Waterpipe tobacco smoking: an emerging health crisis in the United States. Am J Health Behav 2010 Jun;34(3):275-285 [FREE Full text] [Medline: [20001185](https://pubmed.ncbi.nlm.nih.gov/20001185/)]
27. Smith JR, Edland SD, Novotny TE, Hofstetter CR, White MM, Lindsay SP, et al. Increasing hookah use in California. Am J Public Health 2011 Oct;101(10):1876-1879. [doi: [10.2105/AJPH.2011.300196](https://doi.org/10.2105/AJPH.2011.300196)] [Medline: [21852640](https://pubmed.ncbi.nlm.nih.gov/21852640/)]
28. Grekin ER, Ayna D. Waterpipe smoking among college students in the United States: a review of the literature. J Am Coll Health 2012 Apr;60(3):244-249. [doi: [10.1080/07448481.2011.589419](https://doi.org/10.1080/07448481.2011.589419)] [Medline: [22420702](https://pubmed.ncbi.nlm.nih.gov/22420702/)]
29. Cobb CO, Shihadeh A, Weaver MF, Eissenberg T. Waterpipe tobacco smoking and cigarette smoking: a direct comparison of toxicant exposure and subjective effects. Nicotine Tob Res 2011 Feb;13(2):78-87 [FREE Full text] [doi: [10.1093/ntr/ntq212](https://doi.org/10.1093/ntr/ntq212)] [Medline: [21127030](https://pubmed.ncbi.nlm.nih.gov/21127030/)]
30. Smith SY, Curbow B, Stillman FA. Harm perception of nicotine products in college freshmen. Nicotine Tob Res 2007 Sep;9(9):977-982. [doi: [10.1080/14622200701540796](https://doi.org/10.1080/14622200701540796)] [Medline: [17763115](https://pubmed.ncbi.nlm.nih.gov/17763115/)]
31. Primack BA, Hopkins M, Hallett C, Carroll MV, Zeller M, Dachille K, et al. US health policy related to hookah tobacco smoking. Am J Public Health 2012 Sep;102(9):e47-e51. [doi: [10.2105/AJPH.2012.300838](https://doi.org/10.2105/AJPH.2012.300838)] [Medline: [22827447](https://pubmed.ncbi.nlm.nih.gov/22827447/)]
32. Rice VH. Water pipe smoking among the young: the rebirth of an old tradition. Nurs Clin North Am 2012 Mar;47(1):141-148. [doi: [10.1016/j.cnur.2011.10.011](https://doi.org/10.1016/j.cnur.2011.10.011)] [Medline: [22289404](https://pubmed.ncbi.nlm.nih.gov/22289404/)]
33. Foulds J, Veldheer S, Berg A. Electronic cigarettes (e-cigs): views of aficionados and clinical/public health perspectives. Int J Clin Pract 2011 Oct;65(10):1037-1042. [doi: [10.1111/j.1742-1241.2011.02751.x](https://doi.org/10.1111/j.1742-1241.2011.02751.x)] [Medline: [21801287](https://pubmed.ncbi.nlm.nih.gov/21801287/)]
34. Choi K, Fabian L, Mottey N, Corbett A, Forster J. Young adults' favorable perceptions of snus, dissolvable tobacco products, and electronic cigarettes: findings from a focus group study. Am J Public Health 2012 Nov;102(11):2088-2093. [doi: [10.2105/AJPH.2011.300525](https://doi.org/10.2105/AJPH.2011.300525)] [Medline: [22813086](https://pubmed.ncbi.nlm.nih.gov/22813086/)]
35. Freiberg M. Options for state and local governments to regulate non-cigarette tobacco products. Ann Health Law 2012;21(2):407-455. [Medline: [22606921](https://pubmed.ncbi.nlm.nih.gov/22606921/)]
36. 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/)]
37. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. New York: Springer; 2009.
38. Yang Y, Pedersen J. A comparative study on feature selection in text categorization. In: Proceedings of the Fourteenth Annual Conference on Machine Learning.: Morgan Kaufmann Publishers; 1997 Presented at: Fourteenth Annual Conference on Machine Learning; July 8-12, 1997; Nashville, TN p. 412.
39. Forman G. An extensive empirical study of feature selection metrics for text classification. The Journal of Machine Learning Research 2003;3:1289-1305.
40. Rainbow Toolkit. URL: <http://www.cs.cmu.edu/~mccallum/bow/> [accessed 2012-10-13] [WebCite Cache ID 6BOd6Dc7L]
41. Porter Stemmer. URL: <http://tartarus.org/martin/PorterStemmer/> [accessed 2012-10-13] [WebCite Cache ID 6BOdSNgLe]
42. SMART stoplist: list/english. stop URL: <http://jmlr.csail.mit.edu/papers/volume5/lewis04a/a11-smart-stop-list/english.stop> [accessed 2012-10-13] [WebCite Cache ID 6BOdAulJw]

Abbreviations

- API:** Application Programming Interface
- CDC:** Centers for Disease Control
- KNN:** k-nearest neighbors
- SVM:** support vector machine
- USFDA:** United States Federal Drug Administration

Edited by G Eysenbach; submitted 17.01.13; peer-reviewed by N Cobb, S Emery, T Hernández; comments to author 26.03.13; revised version received 05.06.13; accepted 17.06.13; published 29.08.13.

Please cite as:

Myslín M, Zhu SH, Chapman W, Conway M
Using Twitter to Examine Smoking Behavior and Perceptions of Emerging Tobacco Products
J Med Internet Res 2013;15(8):e174
URL: <http://www.jmir.org/2013/8/e174/>
doi: [10.2196/jmir.2534](https://doi.org/10.2196/jmir.2534)
PMID: [23989137](https://pubmed.ncbi.nlm.nih.gov/23989137/)

©Mark Myslín, Shu-Hong Zhu, Wendy Chapman, Mike Conway. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.08.2013. 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

Patient and Public Views on Electronic Health Records and Their Uses in the United Kingdom: Cross-Sectional Survey

Serena A Luchenski¹, BSc, MSc, DFPH; Julie E Reed¹, MChem, PhD, DIC; Cicely Marston², BA (Hons) Oxon, MSc, PhD; Chrysanthi Papoutsis¹, BSc, MSc, DPhil (Oxon); Azeem Majeed³, MB ChB, MD, FRCGP, FFPH; Derek Bell¹, BSc (Hons), MB ChB, MD, FRCP

¹NIHR CLAHRC for Northwest London, Imperial College London, Chelsea & Westminster Hospital NHS Foundation Trust, London, United Kingdom

²Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, United Kingdom

³Department of Primary Care & Public Health, Imperial College London, London, United Kingdom

Corresponding Author:

Derek Bell, BSc (Hons), MB ChB, MD, FRCP

NIHR CLAHRC for Northwest London, Imperial College London, Chelsea & Westminster Hospital NHS Foundation Trust
4th Floor, Lift Bank D, 369 Fulham Road

London, SW10 9NH

United Kingdom

Phone: 44 02033158144

Fax: 44 02087468887

Email: d.bell@imperial.ac.uk

Abstract

Background: The development and implementation of electronic health records (EHRs) remains an international challenge. Better understanding of patient and public attitudes and the factors that influence overall levels of support toward EHRs is needed to inform policy.

Objective: To explore patient and public attitudes toward integrated EHRs used simultaneously for health care provision, planning and policy, and health research.

Methods: Cross-sectional questionnaire survey administered to patients and members of the public who were recruited from a stratified cluster random sample of 8 outpatient clinics of a major teaching hospital and 8 general practices in London (United Kingdom).

Results: 5331 patients and members of the public responded to the survey, with 2857 providing complete data for the analysis presented here. There were moderately high levels of support for integrated EHRs used simultaneously for health care provision, planning and policy, and health research (1785/2857, 62.47%), while 27.93% (798/2857) of participants reported being undecided about whether or not they would support EHR use. There were higher levels of support for specific uses of EHRs. Most participants were in favor of EHRs for personal health care provision (2563/2857, 89.71%), with 66.75% (1907/2857) stating that they would prefer their complete, rather than limited, medical history to be included. Of those “undecided” about integrated EHRs, 87.2% (696/798) were nevertheless in favor of sharing their full (373/798, 46.7%) or limited (323/798, 40.5%) records for health provision purposes. There were similar high levels of support for use of EHRs in health services policy and planning (2274/2857, 79.59%) and research (2325/2857, 81.38%), although 59.75% (1707/2857) and 67.10% (1917/2857) of respondents respectively would prefer their personal identifiers to be removed. Multivariable analysis showed levels of overall support for EHRs decreasing with age. Respondents self-identifying as Black British were more likely to report being undecided or unsupportive of national EHRs. Frequent health services users were more likely to report being supportive than undecided.

Conclusions: Despite previous difficulties with National Health Service (NHS) technology projects, patients and the public generally support the development of integrated EHRs for health care provision, planning and policy, and health research. This support, however, varies between social groups and is not unqualified; relevant safeguards must be in place and patients should be guided in their decision-making process, including increased awareness about the benefits of EHRs for secondary uses.

(*J Med Internet Res* 2013;15(8):e160) doi:[10.2196/jmir.2701](https://doi.org/10.2196/jmir.2701)

KEYWORDS

electronic health records; patient attitudes; health care delivery; research; policy

Introduction

Electronic health records (EHRs) are often heralded as the cornerstone of modern health care provision, although their development and implementation still remains an international challenge [1-6]. In recent years in the United Kingdom, there have been several policy initiatives aiming to alter the technological landscape in the National Health Service (NHS). The initial focus on centralized, top-down national databases, promoted by the National Programme for Information Technology (NPfIT), has now been displaced by the most recent information strategy launched in 2012, "The Power of Information: Putting all of us in control of the health and care information we need" [7]. This document emphasizes information sharing to ensure local EHR systems work seamlessly "across the entire health and social care sector, both within and between organizations" to provide data to multiple stakeholders [7]. In line with this strategic vision, the Department of Health has announced that hospitals should have implemented electronic patient records by 2015, with fully digitized health records being deployed by 2018 across the health care sector [8-10]. In supporting these aims, the Information Governance Review, newly published at the time of writing, recognizes that the duty to share information in the patients' interests can be as important as the duty to ensure confidentiality, although the recommendations do not extend the use of identifiable data [11].

Within the policy arena, patients and members of the public are often presented as the primary beneficiaries of this technologically-orientated agenda [8-10]. However, their attitudes towards sharing medical information have been studied in a fragmented fashion. The larger part of previous research has focused either on specific EHR systems (eg, Summary Care Record [12]) or on the use of segregated data for specific purposes (eg, research [13] or care improvement [14]). Most people are generally in favor of EHRs and information sharing, but differences exist depending on the intended use, the type of information being shared, and whether health information is anonymized or not [12-22]. As such, public support is not unqualified. A range of concerns have been documented, including privacy, security, control over access, use, and potential misuse of data [12,17,23-25]. Previous research further shows differences in opinion by age, education level, socioeconomic situation and health status [16,17,19,26]. Furthermore, those with long-term conditions appear more supportive of EHRs for personal health benefit as well as for research [12,23,26].

As we progress toward implementing the information strategy, we require a more in-depth understanding of attitudes toward EHRs and the factors that influence overall levels of support. Information flows in health care are often complex and data are used for multiple purposes, as for example at the interface of care and research [27]. For this reason, we should assess patient views about EHRs that acknowledge their use for multiple purposes including health care provision, health services policy and planning, as well as research. Previous research has provided only basic information on sociodemographic variables, and there has been little work on associations between attitudes to

EHRs and the experience of patients in health care. People in regular contact with different health services may have encountered difficulties with information sharing between professionals and thus might perceive EHRs as a solution to these communication barriers.

Against this background of policy change within the United Kingdom, this paper surveys patient and public attitudes based on a more complex view of EHRs as systems that may be used for multiple purposes, as well as examining how attitudes differ when considering specific uses, including health care provision, policy and planning, as well as research. The aim of this study is to enhance understanding of patient and public views about the development of universal patient EHRs and their willingness to share their personal records in a national EHR system, by addressing the following questions:

1. What is the level of patient and public support for a national EHR system overall and for what purposes should it be used?
2. What is the relationship between overall support for a national EHR system and the use of EHRs for health care, planning and policy, and health research?
3. How are health, health care use, and sociodemographic characteristics associated with patient and public support for a national EHR system?

Methods

We conducted a cross-sectional self-complete questionnaire survey using a stratified cluster random sample of patients and members of the public in an area of West London, United Kingdom. Participants were recruited in 8 outpatient waiting areas of a university teaching hospital and the waiting rooms of 8 general practice (GP) surgeries within the hospital catchment area over a 6-week period beginning August 1, 2011. Eligibility criteria for participation were: (1) 18 years or older, (2) first time filling in the survey, and (3) able to understand the information describing the research study. In total, 5331 individuals participated in the survey. Full details of the study protocol are published elsewhere [28]. The study was approved by the London Dulwich Research Ethics Committee (Ref No 10/H0808/96).

Data were collected on patient and public views about a national EHR system and the purposes for which EHRs should be used if such a system existed. The front page of the questionnaire introduced participants to EHRs using the following definition: "If created, your electronic health record would store everything about your health and the health care you receive from your birth until your death. Electronic health records would bring together in one record all of your separate files, whether stored on paper or on a computer, in all of the different locations where you get health care." The questionnaire made clear that the study concerned detailed EHRs rather than Summary Care Records. The 31-item questionnaire examined various aspects of patient and public views, but here we present the findings relating to the following 4 key questions:

1. If there was a national electronic health records system, would you want your record to be part of it for your own

- health care? (“Yes, complete record”; “Yes, partial record”; “No”)
- If there was a national electronic health records system, would you want your record to be part of it for health services planning and policy? (“Yes, name and address present”; “Yes, name and address removed”; “No”)
 - If there was a national electronic health records system, would you want your record to be part of it for health research? (“Yes, name and address present”; “Yes, name and address removed”; “No”)
 - Overall, are you in favor of the development of a national electronic health records system? (“Yes”; “No”; “Undecided”)

Further questions recorded details of respondents’ health (whether respondent had a long-term condition or not), health care use (personal health care visits in the previous 6 months) and sociodemographic characteristics (birth year, sex, ethnicity, highest education level attained). The full survey instrument is included in [Multimedia Appendix 1](#).

Only respondents providing complete data for the variables of interest were included in the final statistical analysis (N=2857). We first described the study variables including the number and proportion of the analysis sample. To assess the effects of excluding individuals with missing data, we used logistic regression to compare the distribution of responses for each variable between the analysis sample and the missing sample. We examined the proportions of missing data for questions on the final page of the questionnaire compared with questions at the beginning of the questionnaire to assess the effect of questionnaire length on question completion.

We used descriptive analysis to examine our first 2 questions. The proportions of respondents who would support the development of a national EHR system in the United Kingdom and the proportions of respondents who would allow their EHR to be used for their personal health care, health services planning and policy, and health research were calculated. We then examined the correlation between overall support for a national EHR system and views about the three proposed uses of EHRs using chi-square to test for statistical significance.

We also used a multivariable multinomial regression model to examine associations between views about a national EHR system and health, health care use, and sociodemographic characteristics. We tested for multicollinearity between the independent variables and found all VIF (variance inflation factor) scores to be approximately 1, indicating that they were not highly correlated and could thus be combined in multivariable analyses. *P* values and 95% confidence intervals were adjusted for the clustered design of the survey. All analyses were conducted using Stata IC version 9.0.

Results

Participants

We recruited 5331 respondents representing 85.50% of all individuals approached. In total, 2857 out of 5331 (53.59%) respondents completed all relevant sections of the questionnaire and were included in the final analysis. There was no significant

difference in the rate of completion for questions at the beginning of the questionnaire compared with those at the end, indicating that respondents were able to complete the questionnaire in the time available.

Study Population

The sociodemographic, health, and health care use characteristics of the sample are shown in [Table 1](#). The sample is relatively young, with a high proportion of women, and with a high level of educational attainment, while it is also ethnically diverse. A larger proportion of respondents were sampled in outpatient clinics rather than in GP surgeries, which is a characteristic of the survey design. Hospital outpatient clinics were busier than GP surgeries and patients attending the hospital had a higher proportion of health problems than those routinely attending GP surgeries. The recruitment time was divided equally between the two settings to ensure that individuals with long-term health conditions participated in the survey. The majority of respondents have at least 1 long-term condition and accordingly the sample population are moderately frequent health care users.

Support for a National EHR System and for What Purposes

Respondents’ overall level of support for a national EHR system and the use of EHRs for health care, planning and policy, and health research are presented in [Figure 1](#).

When asked to consider the development of a national EHR system (that would simultaneously support health care, planning and policy, and research), 1785 out of 2857 respondents reported overall support (62.47%), while a large minority of people reported being undecided in their views (n=798, 27.93%). A smaller proportion (n=274, 9.59%) said they would not support a national EHR system used for multiple purposes.

In terms of personal health care provision, responses were more positive with a striking proportion supporting the development of EHRs for this specific purpose (2563 out of 2857, 89.71%). Although 66.75% (n=1907) of respondents would support the use of their complete medical history, almost a quarter of participants (n=656) would allow only limited health information to be part of a national EHR system. 294 out of 2857 (10.29%) said they were opposed to the use of EHRs for health care purposes.

A significant proportion of respondents supported the use of EHRs for planning and policy (n=2274, 79.59%). However, the majority reported that they would only allow their records to be included in an integrated EHR system if personal identifiers had been removed (n=1707, 59.75%). Just one-fifth (n=567, 19.8%) supported the use of identifiable data, with a similar proportion (n=583, 20.4%) opposed to any use of their EHRs for planning and policy.

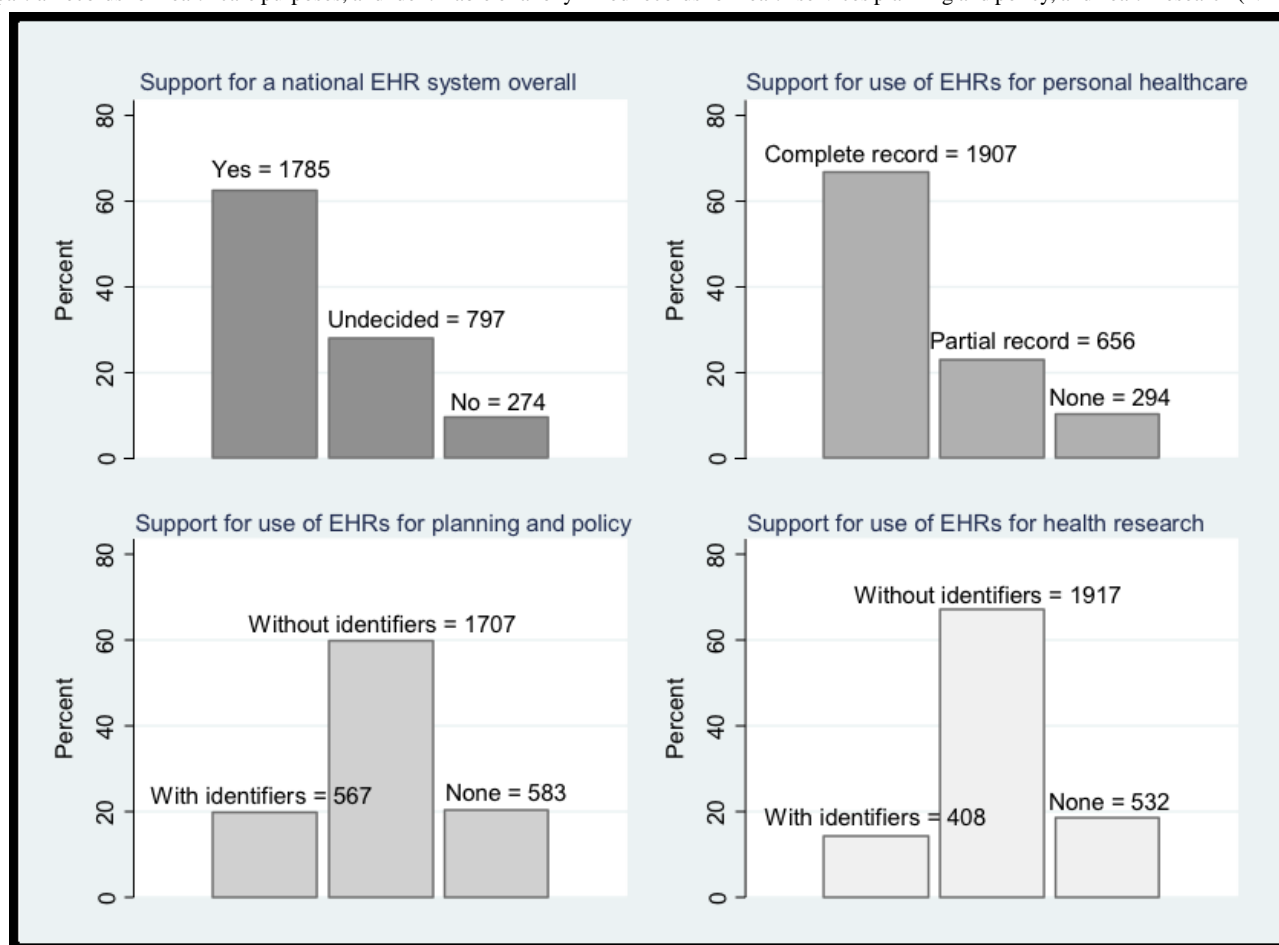
With regard to using national EHRs for health research, 2325 out of 2857 participants would be similarly supportive of having their records included in the system (81.38%). Yet, only 408 (14.28%) of respondents answered that they would allow their identifiable records to be shared, while 1917 (67.10%) of respondents would prefer having their name and address

removed. Almost one-fifth (n=532, 18.62%) said they would specifically not wish their record to be used at all for health research

Table 1. Summary statistics of final analysis sample by sociodemographics, health, and health care use characteristics (N=2857).

Variable	n (%)
Age category	
18-24	226 (7.91)
25-34 (base)	757 (26.50)
35-44	614 (21.49)
45-54	444 (15.54)
55-64	334 (11.69)
65-74	294 (10.29)
75+	188 (6.58)
Sex	
Female (base)	1700 (59.50)
Male	1157 (40.50)
Ethnicity	
White British (base)	1602 (56.07)
White Non-British	583 (20.41)
Black British	207 (7.25)
Asian British	229 (8.02)
Mixed	93 (3.26)
Other	143 (5.01)
Educational qualifications	
None	145 (5.08)
GCSE	319 (11.17)
A-Level	288 (10.08)
Vocational Qualification	335 (11.73)
Degree	1062 (37.17)
Higher Degree (base)	708 (24.78)
Clinic type	
GP (base)	953 (33.36)
Outpatient	1904 (66.64)
Number of health care visits in the past 6m	
0-2 visits (base)	1041 (36.44)
3-5 visits	998 (34.93)
6-9 visits	459 (16.07)
10 plus visits	359 (12.57)
Long-term conditions	
None (base)	1007 (35.25)
At least one condition	1850 (64.75)

Figure 1. Respondents' overall preferences for the development of a national electronic health records system and their views on the use of complete or partial records for health care purposes, and identifiable or anonymized records for health services planning and policy, and health research (N=2857).



Relationship Between Overall Support and Support for Specific Purposes

The relationship between individuals' expressed level of support for a national EHR system and their views about using EHRs for the specific purposes of personal health care, planning and policy, and research are shown in Table 2.

For the 798 (27.93%) of respondents undecided about supporting a national EHR system, the majority (n=696, 87.2%) report that they would support the use of EHRs for their own health care, with 373 out of 798 (46.7%) favoring the use of their complete records and 323 (40.5%) supporting the use of records with limited health information.

Approximately two-thirds of those undecided (n=798) about their overall support for EHRs would support their use for planning and policy (461, 57.77%), and for health research purposes (538, 67.42%), provided the records did not contain personal identifiers.

The majority of those who responded positively (n=1785) to the development of a national EHR system said they would allow their records to be used for health care (1752, 98.15%), planning and policy (1616, 90.53%), and health research (1617, 90.58%). Of those who said they would not be in favor of a national EHR system (n=274), around 40% reported that they would support using EHRs for specific purposes (115 for health care, 101 for planning and policy, and 108 for health research).

Associations Between Overall Support and Sociodemographics, Health, and Health Care Use

Associations between respondents' overall level of support for a national EHR system and their sociodemographics, health, and health care use characteristics are shown in Table 3. This multinomial multivariable analysis is interpreted by comparing those who are undecided to those who would support a national EHR system, as well as comparing those who would not be supportive of EHRs to those who expressed positive attitudes. In effect, it is similar to interpreting 2 separate logistic regression models.

There was no clear pattern of association between age and being undecided on support for EHRs overall, or between age and being supportive of such a system. However, there was a graded association between age and lack of support for a national EHR system with older people increasingly more likely to report that they would not be in favor of such a system compared with 25-34 year olds (the largest age category in the sample).

Men were less likely than women to report that they were undecided compared with being positive about EHRs (RR=0.68, 0.59-0.79). Black and Asian British respondents were also more likely to say that they were undecided in their views on EHRs than to say that they would be supportive compared to White British respondents (RR=1.96, 1.34-2.86). Black British respondents were more likely to say they would not support the

development of a national EHR system compared with White British respondents (RR=3.72, 2.33-5.94).

Respondents with fewer or no academic qualifications are more likely to report being undecided about their attitudes to EHRs than to report being supportive, compared with those with a higher degree. There are no statistically significant educational differences between people who would support the development of national EHRs and those who would not. There were also no significant differences in this respect between those recruited in GP surgeries and those recruited in hospital outpatient clinics. However, respondents from GP surgeries are more likely to report that they were undecided than positive about national EHRs, compared with those who completed the survey as outpatients in the hospital (RR=1.21, 1.08-1.36).

Individuals who use health services more frequently were less likely to report being undecided about EHRs than to answer that they would be in favor of such a system, compared with less frequent users of health services (0-2 times in the past 6 months). The association is statistically significant for very regular users of health care services (10 or more times in the past 6 months) (RR=0.69, 0.60-0.79). Having a long-term

condition was not associated with respondents' views about a national EHR system.

Missing Data Analysis

The analysis of missing data in Table 4 shows that those included in the sample have the same age and sex distribution of those not included in the sample. However, respondents with missing data are significantly more likely to be Black ($P<.001$) or Asian ($P=.02$) than White British. Those with lower education levels are also more likely to have missing data than those with a higher degree. The analysis of missing data also shows that the clinical setting did not affect respondents' likelihood of providing complete data. However, those who have missing data are significantly more likely to use health care services more often and to report no long-term health problems.

Approximately 10% of respondents had missing data on their views about EHRs (ranging from 9.4% to 11.2%), which is lower than for the other analysis variables. However, the analysis showed that those who were excluded from the final analysis sample were significantly more likely to have favorable views towards EHRs for all 4 outcome variables than those who were included ($P<.001$).

Table 2. Relationship between overall support for a national EHR system and views about the use of EHRs for personal health care, health services planning and policy, and health research, with Chi square (χ^2) tests used to test for statistical significance (N=2857).

	Support for a national EHR system			Total n (%)	χ^2	P
	Yes n (%)	Undecided n (%)	No n (%)			
Personal health care						
Complete record	1484 (83.14)	373 (46.74)	50 (18.25)	1907 (66.75)		
Partial record	268 (15.01)	323 (40.48)	65 (23.72)	656 (22.96)		
Neither record	33 (1.84)	102 (12.78)	159 (58.03)	294 (10.29)		
Total	1785 (100.00)	798 (100.00)	274 (100.00)	2857 (100.00)	1107	<.001
Health services planning and policy						
With identifiers	451 (25.27)	96 (12.03)	20 (7.30)	567 (19.85)		
Without identifiers	1165 (65.27)	461 (57.77)	81 (29.56)	1707 (59.75)		
Neither record	169 (9.47)	241 (30.20)	173 (63.14)	583 (20.41)		
Total	1785 (100.00)	798 (100.00)	274 (100.00)	2857 (100.00)	511	<.001
Health research						
With identifiers	338 (18.94)	62 (7.77)	8 (2.92)	408 (14.28)		
Without identifiers	1279 (71.65)	538 (67.42)	100 (36.50)	1917 (67.10)		
Neither record	168 (9.41)	198 (24.81)	166 (60.58)	532 (18.62)		
Total	1785 (100.00)	798 (100.00)	274 (100.00)	2857 (100.00)	467	<.001

Table 3. Relative risks (RR) indicating associations between overall support for a national EHR system and sociodemographic, health, and health care use characteristics. Multinomial logistic regression model comparing those that would support the development of EHRs overall (base=Yes), compared with those who are undecided and those who would not support EHRs; *P* values and 95% CI adjusted for clustering by sampling site (N=2857).

Respondent characteristics	Overall views on the development of a national EHR system ^a (base: In favor)					
	Undecided			Against		
	Adjusted RR	95% CI	<i>P</i> value	Adjusted RR	95% CI	<i>P</i> value
Age (base: 25-34)						
18-24	1.59	(1.13, 2.24)	.008	1.56	(0.83, 2.92)	.17
35-44	1.02	(0.79, 1.31)	.90	1.66	(1.17, 2.34)	.004
45-54	1.19	(0.94, 1.51)	.14	2.29	(1.39, 3.77)	<.001
55-64	1.49	(1.09, 2.03)	.01	2.60	(1.70, 3.98)	<.001
65-74	0.84	(0.58, 1.23)	.37	2.53	(1.51, 4.22)	<.001
75+	0.97	(0.65, 1.46)	.89	2.86	(1.83, 4.47)	<.001
Sex (base: female)						
Male	0.68	(0.59, 0.79)	<.001	0.88	(0.67, 1.15)	.36
Ethnicity (base: White British)						
White non-British	1.14	(0.93, 1.40)	.22	1.00	(0.75, 1.32)	.98
Black British	1.96	(1.34, 2.86)	<.001	3.72	(2.33, 5.94)	<.001
Asian British	1.43	(1.03, 1.99)	.03	1.37	(0.88, 2.14)	.17
Mixed	1.40	(0.97, 2.04)	.08	1.07	(0.55, 2.09)	.85
Other	1.23	(0.80, 1.90)	.35	1.18	(0.79, 1.78)	.42
Education (base: higher degree)						
None	1.58	(1.03, 2.44)	.04	1.25	(0.60, 2.57)	.55
GCSE	1.96	(1.40, 2.75)	<.001	1.27	(0.75, 2.16)	.38
A-Level	1.51	(1.08, 2.10)	.02	1.00	(0.56, 1.77)	1.00
Vocational	1.51	(1.20, 1.90)	<.001	0.85	(0.47, 1.55)	.59
Degree	1.29	(1.05, 1.59)	.02	0.93	(0.76, 1.14)	.48
Clinic type (base: GP clinic)						
Outpatient clinic	1.21	(1.08, 1.36)	<.001	1.13	(0.86, 1.48)	.38
Number of health care visits in the past 6 months (base: 0-2 visits)						
3-5 visits	0.93	(0.76, 1.15)	.51	0.80	(0.60, 1.05)	.11
6-9 visits	0.86	(0.67, 1.09)	.21	0.67	(0.40, 1.12)	.13
10 plus visits	0.69	(0.60, 0.79)	<.001	1.21	(0.71, 2.06)	.49
Reports long-term medical conditions (base: no conditions)						
At least 1 condition	1.21	(0.92, 1.58)	.17	1.35	(0.93, 1.95)	.11

^aThe questionnaire asked: Overall, are you in favor of the development of a national electronic health records system? (“Yes”; “No”; “Undecided”).

Table 4. Support for EHR: Univariable logistic regression of missing data by respondent characteristics (N=5331).

Variable	Missing (%)	Odds ratio	95% CI	P
Age category (base: 25-34)	799 (14.99)			
18-24		0.87	(0.67, 1.15)	.33
35-44		0.95	(0.81, 1.13)	.56
45-54		0.90	(0.71, 1.13)	.36
55-64		0.95	(0.81, 1.13)	.58
65-74		0.97	(0.77, 1.21)	.78
75+		0.87	(0.68, 1.11)	.25
Sex (base: female)	611 (11.46)			
Male		1.10	(0.96, 1.27)	.17
Ethnicity (base: White British)	1109 (20.80)			
White non-British		1.14	(1.00, 1.31)	.047
Black British		0.62	(0.51, 0.75)	<.001
Asian British		0.71	(0.53, 0.94)	.02
Mixed		0.86	(0.68, 1.10)	.23
Other		0.77	(0.56, 1.05)	.10
Educational qualifications (base: higher degree)	833 (15.63)			
None		0.61	(0.49, 0.75)	<.001
GCSE		0.90	(0.78, 1.03)	.13
A-Level		0.76	(0.61, 0.94)	.01
Vocational qualification		0.87	(0.65, 1.16)	.33
Degree		0.91	(0.78, 1.06)	.21
Clinic type (base: GP)	0 (0)			
Outpatient		1.01	(0.86, 1.20)	.87
Number of health care visits in the past 6m (base: 0-2 visits)	686 (12.87)			
3-5 visits		1.16	(1.05, 1.29)	.003
6-9 visits		1.48	(1.27, 1.74)	<.001
10 plus visits		1.17	(0.97, 1.41)	.10
Long-term conditions (base: none)	1103 (20.69)			
At least 1 condition		0.68	(0.58, 0.81)	<.001
Overall support for EHRs (base: yes)	584 (10.95)			
Undecided		0.67	(0.59, 0.74)	<.001
No		0.43	(0.36, 0.50)	<.001
Support for EHRs used for health care purposes (base: complete record)	499 (9.36)			
Partial record		0.71	(0.59, 0.84)	<.001
No		0.43	(0.37, 0.51)	<.001
Support for EHRs used for health services planning and policy purposes (base: without identifiers)	566 (10.62)			

Variable	Missing (%)	Odds ratio	95% CI	P
With identifiers		0.78	(0.68, 0.90)	.001
No		0.56	(0.49, 0.65)	<.001
Support for EHRs used for health research purposes (base: without identifiers)	599 (11.24)			
With identifiers		0.78	(0.68, 0.89)	<.001
No		0.50	(0.44, 0.58)	<.001

Discussion

Principal Findings

This study suggests that there is general support for the development of a national EHR system that would simultaneously use data for multiple purposes, such as personal health care, policy and planning, as well as health research. However, an important minority—about a quarter of participants (n=798, 27.93%)—remain undecided in their views, and nearly 10% (n=274) would be opposed to such a system. When asked about specific purposes for EHRs, over two-thirds of all respondents would support the inclusion of their full medical history and personally identifiable information for personal health care provision. In contrast, for health policy, planning, and research uses, higher support was expressed for use of anonymized EHRs. Even in the group expressing overall negative views towards an integrated EHR system (n=274, 9.59%), there are respondents who would still choose to participate in EHRs if their information was used for specific purposes, such as for their personal health care (n=115, 42.0%), policy and planning (n=101, 36.9%), or health research (n=108, 39.4%). Similarly, over 86% of those undecided (696 out of 798) in their level of support for a national EHR system are supportive of full or partial records being used specifically for their personal health care.

This study also shows significant differences in levels of support depending on sociodemographic characteristics. Age appears to play an important role in support for EHRs with older participants significantly less in favor of EHRs than younger respondents. Black British respondents also show significantly less support than respondents of other ethnic groups. In addition, educational attainment and patterns of health care use differentiate those who report being undecided in their views on EHRs from those who answer that they would be in favor of a national EHR system. However, there is no association between having a long-term condition as measured in this study and support for a national EHR system.

Strengths and Limitations

This is the first large study to explore patient and public attitudes towards EHRs in the United Kingdom and also the first to draw on a more complex and comprehensive picture of the different potential uses of EHRs, rather than examining their use only for specific purposes. To minimize selection bias, we recruited participants at different days and times following a random sampling design. Although the overall response rate was very high (85.50%), only half of the participants completed the

questions for the variables analyzed in this paper (2857 out of 5331, 53.59%). The analysis of missing data shows that there are no age or sex differences between those who were included in the final analysis sample and those who were excluded, but there were ethnicity and education differences. Notably, those with less favorable views were more likely to be excluded from the final analysis. In terms of confounding factors, we measured and adjusted for the main confounding variables in our multivariable analysis; however, the results could still be affected by unmeasured confounders, such as overall levels of trust in the government and authorities. Other methodological considerations related to possible sources of measurement bias are discussed in [28].

Previous Studies

While other quantitative and qualitative studies have reported that patients and the public would generally support EHRs [12,16] our results contradict previous studies in the United Kingdom and Ireland, which have found higher levels of support in older age groups for information sharing in medical research or GP summary records [17,29]. However, our findings are consistent with similarly large studies in other countries that have found older age groups to be less supportive of EHRs [16]. Our study resonates with previous research showing that ethnic background affects attitudes towards health information sharing: people from BME (Black and Minority Ethnic) communities or people who do not identify themselves as White British have been shown to be less inclined to allow their data to be used for public health and medical research [19,30,31]. In addition, our results on educational differences in opinion between being undecided and being in favor of integrated EHRs extends previous work showing that higher levels of educational attainment are associated with willingness to share health information and support for EHRs [17,29]. Recruitment was carried out in the outpatient and GP population of West London, United Kingdom. Respondents were ethnically diverse with a spectrum of educational backgrounds, which allowed us to sample opinions from a wide range of sociodemographic groups. Overall patterns of opinions may be similar in other areas of London given similarities in sociodemographic and health care characteristics.

Implications for Research and Policy

The study shows that a proportion of people currently unsupportive or undecided about national EHRs for multiple purposes may nevertheless be amenable to EHRs being used for clearly defined purposes. Patient and public perceptions about inclusion of their records in EHRs for their personal health

care mirror levels of overall support for national EHRs, suggesting that considerations of personal health needs might be driving these opinions. Additionally, sociodemographic disparities in levels of support indicate that preferences cannot be considered homogeneous. Introduction of national EHRs may risk widening inequalities for BME groups and the elderly, who are more likely to be against the development of this system. Wider sharing of information may have an effect on their trust toward the health care system and their willingness to seek medical help. Less information on conditions affecting BME and elderly groups may also impact negatively on the potential for health research relevant to these populations and on the planning for services to support their needs. More in-depth research on patient views is needed to draw out the nuances involved in decision-making processes related to wider sharing of health information. Qualitative research studies will enhance our understanding in this area. A more nuanced understanding also has practical relevance in terms of framing policy messages when an EHR is launched and publicized; gaining the support of undecided or opposed groups as well as the public in general could determine whether or not EHRs can be successfully implemented as planned.

Given the well-documented problems inherent in current systems for exchanging patient information between health care professionals and organizations, we hypothesized that the respondents with greater levels of exposure to the health care system would be more acutely aware of the limitations of the current systems and therefore show greater levels of support for EHRs. However, our results in this paper have not indicated

a clear relationship between personal health or health care experience and levels of support for EHRs. This suggests that we need to consider how or whether the nuances of health care experience might affect levels of support and use of EHR systems. Understanding an individual's broader relationship with health care including the need to visit different types of health services, and levels of trust and satisfaction with previous health care encounters may provide greater insight in to the relationship between individuals and their support for EHRs.

Conclusions

Despite the limited success of the NPfIT program in the United Kingdom, there are high levels of support among patient populations for the establishment of national EHRs. Levels of support are not homogenous and the perspectives of the elderly and Black British populations in particular need to be understood more thoroughly to ensure EHRs do not contribute to widening inequalities in health.

Support is greatest for use of EHRs for personal health care. While support for policy and planning and research is also high, most respondents preferred partial or anonymous data to be used for information sharing rather than complete health records. Our results also suggest that individuals who are currently opposed to, or undecided about the introduction of EHRs for multiple purposes, are nevertheless more likely to be supportive if specific conditions are met regarding the content and purpose of EHRs. Such knowledge can help inform the provision of information for and engagement with specific patient and public groups to ensure that the design of any EHR system is acceptable and effective.

Acknowledgments

This paper presents independent research funded by the Wellcome Trust [086124/Z/08/Z]. The Wellcome Trust had no role in the study design; the collection, analysis, and interpretation of data; the writing of the paper; or the decision to submit it for publication. JR and DB are supported by NIHR CLAHRC for Northwest London, and JR is also supported by the Health Foundation, an independent charity working to continuously improve the quality of health care in the United Kingdom. This paper presents independent research commissioned by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) program for North West London. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

We would like to thank the following contributions to this study: Kaori Sasaki contributed to the design and coordination of the study. Anjali Balasanthiran coordinated the acquisition of data with support from Rachael Aldersley, Cameron Bell, Sylvia Chalkley, Jason Curran, Shaun D'Souza, Stuart Green, Sarah Hancox, Sina Iqbal, Uzoma Nnajiuba, Harsita Patel, Joshua Wolrich, and Jade Zhao. We are also grateful to our reviewers for their useful comments.

Authors' Contributions

All authors of this study meet the requirements for authorship and have approved the final version of the paper to be published. SL designed the data collection tools and sampling methodology, planned and conducted the data analysis, and drafted and revised the paper. JR conceived of the study, provided oversight to its design and coordination, contributed to the interpretation of the data, and drafted and revised the paper. CM contributed to the design of the study, data collection tools, the interpretation of the data, and revised the paper. CP contributed to the interpretation of the data and drafted and revised the paper. AM contributed to the design and coordination of the study and reviewed the paper. DB conceived of the study, contributed to its design and coordination, and revised the paper. DB is guarantor of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questionnaire.

[[PDF File \(Adobe PDF File\), 561KB - jmir_v15i8e160_app1.PDF](#)]

References

1. Majeed A. Sources, uses, strengths and limitations of data collected in primary care in England. *Health Stat Q* 2004(21):5-14. [Medline: [15615148](#)]
2. Sheikh A, Cornford T, Barber N, Avery A, Takian A, Lichtner V, et al. Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in "early adopter" hospitals. *BMJ* 2011;343:d6054 [[FREE Full text](#)] [Medline: [22006942](#)]
3. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. *BMC Med* 2011;9:46 [[FREE Full text](#)] [doi: [10.1186/1741-7015-9-46](#)] [Medline: [21524315](#)]
4. Hiller J, McMullen MS, Chumney WM, Baumer DL. Privacy and Security in the Implementation of Health Information Technology (Electronic Health Records): US and EU Compared. *BUJ Sci & Tech L* 2011;17:1-39.
5. Coiera E. Do we need a national electronic summary care record? *Med J Aust* 2011 Jan 17;194(2):90-92. [Medline: [21241224](#)]
6. Nazi KM. The personal health record paradox: health care professionals' perspectives and the information ecology of personal health record systems in organizational and clinical settings. *J Med Internet Res* 2013;15(4):e70 [[FREE Full text](#)] [doi: [10.2196/jmir.2443](#)] [Medline: [23557596](#)]
7. Department of Health. The power of information: Putting all of us in control of the healthcare information we need. London, UK; 2012. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/134336/dh_134205.pdf [accessed 2013-06-03] [[WebCite Cache ID 6H6Lu1G6a](#)]
8. Department of Health. Press release: Jeremy Hunt challenges NHS to go paperless by 2018. London, UK; 2013. URL: <https://www.gov.uk/government/news/jeremy-hunt-challenges-nhs-to-go-paperless-by-2018--2> [accessed 2013-06-03] [[WebCite Cache ID 6H6N7aICc](#)]
9. Department of Health. The Mandate - A mandate from the Government to the NHS Commissioning Board: April to March 2012. London, UK; 2013. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/127193/mandate.pdf [accessed 2013-06-03] [[WebCite Cache ID 6H6MckbDd](#)]
10. NHS Commissioning Board. Everyone counts: Planning for Patients 2013/14. London, UK; 2013. URL: <http://www.england.nhs.uk/wp-content/uploads/2013/04/ec-tech-def.pdf> [accessed 2013-06-03] [[WebCite Cache ID 6H6MefY3E](#)]
11. Department of Health. Information: to share or not to share? The Information Governance Review. London, UK; 2013. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf [accessed 2013-06-03] [[WebCite Cache ID 6H6MiAqBp](#)]
12. Greenhalgh T, Wood GW, Bratan T, Stramer K, Hinder S. Patients' attitudes to the summary care record and HealthSpace: qualitative study. *BMJ* 2008 Jun 7;336(7656):1290-1295 [[FREE Full text](#)] [doi: [10.1136/bmj.a114](#)] [Medline: [18511764](#)]
13. Willison DJ, Swinton M, Schwartz L, Abelson J, Charles C, Northrup D, et al. Alternatives to project-specific consent for access to personal information for health research: insights from a public dialogue. *BMC Med Ethics* 2008;9:18 [[FREE Full text](#)] [doi: [10.1186/1472-6939-9-18](#)] [Medline: [19019239](#)]
14. Weitzman ER, Kelemen S, Kaci L, Mandl KD. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. *BMC Med Inform Decis Mak* 2012;12:39 [[FREE Full text](#)] [doi: [10.1186/1472-6947-12-39](#)] [Medline: [22616619](#)]
15. Whiddett R, Hunter I, Engelbrecht J, Handy J. Patients' attitudes towards sharing their health information. *Int J Med Inform* 2006 Jul;75(7):530-541. [doi: [10.1016/j.ijmedinf.2005.08.009](#)] [Medline: [16198142](#)]
16. EKOS Research Associates. Electronic Health Information and Privacy Survey: What Canadians Think: Canada Health Infoway, Health Canada, and the Office of the Privacy Commissioner of Canada. 2007. URL: https://www.infoway-inforoute.ca/index.php/resources/reports/privacy/doc_download/78-ekos-survey-on-electronic-health-information-and-privacy-full [accessed 2013-06-03] [[WebCite Cache ID 6H6MnEhBu](#)]
17. Ipsos MORI. The Use of Personal Health Information in Medical Research: General Public Consultation Final Report: Medical Research Council. 2007. URL: http://www.mrc.ac.uk/consumption/ideplg?IdcService=GET_FILE&dID=10983&dDocName=MRC003810&allowInterrupt=1 [[WebCite Cache ID 6H6MpN5Av](#)]
18. Kass NE, Natowicz MR, Hull SC, Faden RR, Plantinga L, Gostin LO, et al. The use of medical records in research: what do patients want? *J Law Med Ethics* 2003;31(3):429-433. [Medline: [14626550](#)]
19. Barrett G, Cassell JA, Peacock JL, Coleman MP, National Cancer Registry. National survey of British public's views on use of identifiable medical data by the National Cancer Registry. *BMJ* 2006 May 6;332(7549):1068-1072 [[FREE Full text](#)] [doi: [10.1136/bmj.38805.473738.7C](#)] [Medline: [16648132](#)]

20. Department of Health. Summary of Responses to the Consultation on the Additional Uses of Patient Data. London, UK; 2009. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_110715.pdf [accessed 2013-06-03] [WebCite Cache ID 6H6Mvdf5o]
21. Campbell B, Thomson H, Slater J, Coward C, Wyatt K, Sweeney K. Extracting information from hospital records: what patients think about consent. *Qual Saf Health Care* 2007 Dec;16(6):404-408 [FREE Full text] [doi: [10.1136/qshc.2006.020313](https://doi.org/10.1136/qshc.2006.020313)] [Medline: [18055882](https://pubmed.ncbi.nlm.nih.gov/18055882/)]
22. Powell J, Fitton R, Fitton C. Sharing electronic health records: the patient view. *Inform Prim Care* 2006;14(1):55-57. [Medline: [16848967](https://pubmed.ncbi.nlm.nih.gov/16848967/)]
23. Pyper C, Amery J, Watson M, Crook C. Access to electronic health records in primary care—a survey of patients' views. *Med Sci Monit* 2004 Nov;10(11):SR17-SR22. [Medline: [15507869](https://pubmed.ncbi.nlm.nih.gov/15507869/)]
24. Paterson L, Grant L. Privacy and prejudice: Young people's views on the development and use of Electronic Patient Records. London: The Royal Academy of Engineering 2010 URL: http://www.raeng.org.uk/news/publications/list/reports/Privacy_and_Prejudice_EPR_views.pdf [accessed 2013-06-03] [WebCite Cache ID 6H6N1yKeI]
25. Simon SR, Evans JS, Benjamin A, Delano D, Bates DW. Patients' attitudes toward electronic health information exchange: qualitative study. *J Med Internet Res* 2009;11(3):e30 [FREE Full text] [doi: [10.2196/jmir.1164](https://doi.org/10.2196/jmir.1164)] [Medline: [19674960](https://pubmed.ncbi.nlm.nih.gov/19674960/)]
26. Zulman DM, Nazi KM, Turvey CL, Wagner TH, Woods SS, An LC. Patient interest in sharing personal health record information: a web-based survey. *Ann Intern Med* 2011 Dec 20;155(12):805-810. [doi: [10.7326/0003-4819-155-12-201112200-00002](https://doi.org/10.7326/0003-4819-155-12-201112200-00002)] [Medline: [22184687](https://pubmed.ncbi.nlm.nih.gov/22184687/)]
27. Powell J, Buchan I. Electronic health records should support clinical research. *J Med Internet Res* 2005;7(1):e4 [FREE Full text] [doi: [10.2196/jmir.7.1.e4](https://doi.org/10.2196/jmir.7.1.e4)] [Medline: [15829476](https://pubmed.ncbi.nlm.nih.gov/15829476/)]
28. Luchenski S, Balasanthiran A, Marston C, Sasaki K, Majeed A, Bell D, et al. Survey of patient and public perceptions of electronic health records for healthcare, policy and research: study protocol. *BMC Med Inform Decis Mak* 2012;12:40 [FREE Full text] [doi: [10.1186/1472-6947-12-40](https://doi.org/10.1186/1472-6947-12-40)] [Medline: [22621621](https://pubmed.ncbi.nlm.nih.gov/22621621/)]
29. Buckley BS, Murphy AW, MacFarlane AE. Public attitudes to the use in research of personal health information from general practitioners' records: a survey of the Irish general public. *J Med Ethics* 2011 Jan;37(1):50-55. [doi: [10.1136/jme.2010.037903](https://doi.org/10.1136/jme.2010.037903)] [Medline: [21071570](https://pubmed.ncbi.nlm.nih.gov/21071570/)]
30. Shavers VL, Lynch CF, Burmeister LF. Racial differences in factors that influence the willingness to participate in medical research studies. *Ann Epidemiol* 2002 May;12(4):248-256. [Medline: [11988413](https://pubmed.ncbi.nlm.nih.gov/11988413/)]
31. Corbie-Smith G, Thomas SB, St George DM. Distrust, race, and research. *Arch Intern Med* 2002 Nov 25;162(21):2458-2463. [Medline: [12437405](https://pubmed.ncbi.nlm.nih.gov/12437405/)]

Abbreviations

BME: Black and minority ethnic

CLAHRC: Collaborations for Leadership in Applied Health Research and Care

EHR: electronic health record

GP: general practice

NIHR: National Institute for Health Research

NHS: National Health Service

NPfIT: National Programme for IT

RR: relative risk

Edited by G Eysenbach; submitted 04.05.13; peer-reviewed by T Cornford, J Powell, B Tennant, H Perry; comments to author 25.05.13; accepted 08.06.13; published 23.08.13.

Please cite as:

Luchenski SA, Reed JE, Marston C, Papoutsis C, Majeed A, Bell D

Patient and Public Views on Electronic Health Records and Their Uses in the United Kingdom: Cross-Sectional Survey

J Med Internet Res 2013;15(8):e160

URL: <http://www.jmir.org/2013/8/e160/>

doi: [10.2196/jmir.2701](https://doi.org/10.2196/jmir.2701)

PMID: [23975239](https://pubmed.ncbi.nlm.nih.gov/23975239/)

©Serena A Luchenski, Julie E Reed, Cicely Marston, Chrysanthi Papoutsis, Azeem Majeed, Derek Bell. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 23.08.2013. 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

Limited Use of Price and Quality Advertising Among American Hospitals

David B Muhlestein¹, MS, MHA, JD, PhD; Chrisanne E A Wilks¹, MPA; Jason P Richter¹, MS, MBA, MHA, PhD

Division of Health Services Management and Policy, The Ohio State University College of Public Health, Columbus, OH, United States

Corresponding Author:

David B Muhlestein, MS, MHA, JD, PhD
Division of Health Services Management and Policy
The Ohio State University College of Public Health
200D Cunz Hall
1841 Neil Ave
Columbus, OH, 43222
United States
Phone: 1 (614) 754 8586
Fax: 1 (614) 247 7121
Email: muhlestein.1@osu.edu

Abstract

Background: Consumer-directed policies, including health savings accounts, have been proposed and implemented to involve individuals more directly with the cost of their health care. The hope is this will ultimately encourage providers to compete for patients based on price or quality, resulting in lower health care costs and better health outcomes.

Objective: To evaluate American hospital websites to learn whether hospitals advertise directly to consumers using price or quality data.

Methods: Structured review of websites of 10% of American hospitals (N=474) to evaluate whether price or quality information is available to consumers and identify what hospitals advertise about to attract consumers.

Results: On their websites, 1.3% (6/474) of hospitals advertised about price and 19.0% (90/474) had some price information available; 5.7% (27/474) of hospitals advertised about quality outcomes information and 40.9% (194/474) had some quality outcome data available. Price and quality information that was available was limited and of minimal use to compare hospitals. Hospitals were more likely to advertise about service lines (56.5%, 268/474), access (49.6%, 235/474), awards (34.0%, 161/474), and amenities (30.8%, 146/474).

Conclusions: Insufficient information currently exists for consumers to choose hospitals on the basis of price or quality, making current consumer-directed policies unlikely to realize improved quality or lower costs. Consumers may be more interested in information not related to cost or clinical factors when choosing a hospital, so consumer-directed strategies may be better served before choosing a provider, such as when choosing a health plan.

(*J Med Internet Res* 2013;15(8):e185) doi:[10.2196/jmir.2660](https://doi.org/10.2196/jmir.2660)

KEYWORDS

hospitals; patients; quality indicators; commerce

Introduction

Numerous policies intending to improve the health care system in the United States have been proposed and some have been adopted. Some policies, called consumer-directed strategies, have attempted to change the health care system by targeting the behavior of the consumer of services. For instance, in 2003, health savings accounts and high-deductible insurance plans were embraced with the intent to encourage consumers to be

price conscious at the time of service [1,2]. Some states have mandated some transparency in hospital pricing to encourage this behavior [3,4]. These strategies suppose that price-conscious consumer behavior will lead to providers increasing health care value in the form of better quality or lower prices [5]. For the policymaker's consumer-based system to increase value, there are 3 criteria that must be met: (1) consumers must have access to information on costs and quality, (2) they must choose providers based on those factors, and (3) providers must compete with one another to lower costs or improve quality [6].

There is no guarantee, however, that a consumer-directed system will necessarily lead to hospital competition on price or quality. Even if valid measures of price and quality are available, consumers must still choose hospitals based on these factors. Choice of providers, unlike purchasing fungible goods in the marketplace, is dictated by many nonprice factors, including insurance status, physician recommendations, location, institutional perception, and patient experience. When choosing insurance, cost is an important factor, but once the insured party becomes ill, other factors not related to price or quality are likely to dominate, such as proximity to family or prior relationship with a physician. If consumers are able to direct their care but choose providers based on considerations not related to price or quality, a consumer-directed policy will not lead to increased value.

For a consumer-directed system to lead to improved value, hospitals, which represent the largest proportion of American health care spending, must be responsive to needs and choices of consumers [7]. If hospitals currently are responsive to consumer preference for lower prices and higher quality, there will be evidence that they are competing for consumers based on price or quality. If hospitals are not competing on price and quality, but information on these factors is made available, there is the possibility that hospitals could compete on these factors to increase consumer value in the future. This study will identify whether hospitals are currently competing on price and quality and, if not, what factors they are competing on.

Methods

Overview

It is estimated that 78% of Americans currently use the Internet, that 80% of Internet users compare health care options online, and 58% use it to obtain health information [8-10]. Websites are a nearly universally accessible resource that consumers have to compare services. Because of the ease of access and the prevalence of consumers using the Internet to obtain health care information, hospital websites represent a good source of information to determine what information on price and quality is available to consumers and identify which factors hospitals use to compete for individual consumers. This study presents findings from a review of American hospitals' use of advertising on their company websites. No dataset exists which has systematically evaluated the approaches that hospitals use on their websites to advertise to individual patients; therefore, we performed an original evaluation on how hospitals use varying approaches to entice consumers to use a facility.

Data Collection

The sampling frame for this study was all Medicare-registered hospitals as of July 2012 [11]. This represents 4739 hospitals from all 50 states, the District of Columbia, and several United States territories. Hospitals that register with Medicare include acute care, critical access, and children's hospitals as well as hospitals administered by the Department of Veterans Affairs (VA). We used a 10% simple random sample of these hospitals, resulting in a sample size of 474 hospitals. Information on the hospitals included in our sample and on our sampling frame can be found in Table 1. Information on bed size was obtained

by matching hospitals to Centers for Medicare and Medicaid Services (CMS) Healthcare Cost Report Information System (HCRIS) data. For sampled hospitals, we researched hospital websites for missing values (VA hospitals do not submit CMS cost reports). For the total sample, missing bed sizes were excluded. A hospital was considered urban if it was located in a metropolitan area of a core-based statistical area (CBSA), indicating a regional population of at least 50,000 (determined using a zip code-to-CBSA crosswalk).

The 474 hospitals were divided among the 3 authors and were reviewed between July and September 2012. To insure uniform review of the websites, the reviewers initially evaluated several websites as a group to come to a consensus on evaluating site elements. To test for agreement, multiple websites were reviewed independently by the reviewers; Fleiss' kappa values indicate substantial agreement among all reviewers on identifying categories (kappa=0.633) with slightly lower agreement (kappa=0.571) among specific levels within categories [12]. Any hospitals that did not have a website were confirmed by at least 2 reviewers. Some hospitals that are part of systems do not have dedicated websites, but instead have a webpage as part of the system's website. We used the hospital's webpage for our analysis unless it was so basic as to not mention any services the hospital provides; in these cases we used the system's website (6.1%, 29/474 of hospitals).

Website Evaluation

The focus of our research was on the home page of the website where website visitors are most likely to reach first when researching a hospital. On each reviewed website, the authors sought to identify each of the means in which hospitals may attract customers and refer to these means collectively as forms of advertising. These include formal advertising, such as banners for specific service lines, as well as other content, such as patient education material, that may entice a consumer to use that hospital. Specifically, we sought to identify content related to cost, quality, price, patient safety, customer satisfaction, personal stories, amenities, service lines, access, technology, research, awards, patient education, affiliations, and employment opportunities. Table 2 contains descriptions of these categories of content. Content could fall into more than 1 category, such as an advertisement for 3D mammography representing both technology (the 3D equipment) and a service line (radiology).

After a category of content was identified, it was then classified based on the prominence of the information on the home page as the page's major focus, a minor focus, or a link. The major focus is the primary content in the body of the page, generally including a header and graphical elements or pictures; it is often near the top-center of the page and it is what the eye is usually first drawn to. There can only be 1 major focus on a page, but the major focus can rotate through a number of individual topics, in which case we captured all the categories of content shown on the rotating image. A minor focus is on the home page and includes pictures or content smaller in size compared to the major focus but which contains more than just a link; there can be multiple minor foci on a page. The final category includes text links that navigate the reader to another page. Links may be constantly visible on the website or accessible only through

drop-down menus. Reviewers were also able to add, in a free response section, any comments about the individual categories or the overall website.

In addition to content on the home page, we hypothesized that consumers who were interested in price and quality information

would search on a page to find more information. On sites that had a search feature, we additionally searched for the following terms: cost, price, quality, and patient safety. We then clicked on any link on the first page of results that had the search term in the title or preview and checked the page that was linked to for content relating to price, quality, or patient safety.

Table 1. Summary statistics of the hospitals included in this study.

Characteristic	Sample (N=474)		All US hospitals, % (N=4739)
	n	%	
Hospital is still in operation	472	99.6	
Hospital has a website	453	95.6	
Hospital region			
New England	19	4.0	4.0
Middle Atlantic	41	8.6	9.0
East North Central	82	17.3	15.6
West North Central	57	12.0	13.6
South Atlantic	67	14.1	14.6
East South Central	43	9.1	8.5
West South Central	70	14.8	14.6
Mountain	49	10.3	7.9
Pacific	42	8.9	11.0
Associated Areas	4	0.8	1.2
Urban/rural status			
Urban	277	58.4	60.4
Rural	197	41.6	39.6
Number of beds			
<25	77	16.2	17.3
25-49	104	21.9	22.5
50-99	80	16.9	16.3
100-199	105	22.2	22.6
200-299	56	11.8	10.8
300-399	23	4.9	5.6
400-499	15	3.2	2.3
500+	14	3.0	2.5
Hospital ownership			
Nonprofit	268	56.5	58.0
For profit	92	19.4	16.9
Government	114	24.1	25.0
Hospital type			
Acute care	351	74.1	73.3
Acute care veterans administration	13	2.7	2.7
Children's	1	0.2	0.5
Critical access	109	23.0	23.5

Table 2. Categories of advertising on hospital websites.

Category	Description
Price	Any information on the price or cost of services. This includes information on the price for specific procedures or services as well as information on the average cost of services for a particular diagnosis. The information could deal with the list prices, average reimbursed price, or the expected copayments of the patient. This could be located on the home page or found via a website search.
Quality outcomes	Quantifiable information on outcomes of care. These include data on process or outcome measures, or comparison statistics in which the hospital's outcomes are presented relative to other hospitals. Qualitative descriptions of quality were not included in this category, such as a statement that the hospital is "a regional leader in cardiovascular outcomes," unless accompanied by some quantifiable data. This could be located on the home page, found via search, or linked to offsite. An example of this occurs when a hospital publishes its own scores on 30-day readmissions from Medicare's Hospital Compare website or informs the consumer about data available from an offsite source and links the consumer to the external website.
Patient safety	Quantifiable depiction of patient safety outcomes. This includes information on rates of hospital-acquired infections, pressure ulcers, medication errors, falls, surgical errors, etc. Information could be found on the home page or via search; data could be located on the website or available via a link to an external website.
Customer satisfaction	Data on previous patients' experiences with the hospital. This could include information from Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) surveys, Press Ganey patient satisfaction surveys, or some other numerical or comparative depiction of patient satisfaction with care. An example is the percent of patients that would recommend the hospital to their friends or family.
Awards and accreditation	Formal recognition from an outside entity. Examples include being a top hospital in some specialty or receiving accreditation for a service.
Personal stories	Anecdotal experiences from patients or staff that recount the care they received or provided at the hospital. This includes written experiences and videos relating to how the hospital served the patient or the experience of staff while working at the hospital and serving patients.
Amenities	Includes references to the physical facilities, such as buildings, parking, cafeterias, and gift shops or support services, such as consumer advocates, chaplains, and support groups. Also includes qualitative descriptions of the hospital, such as the atmosphere patients will experience, and nonmedical information available at the hospital, such as cooking classes or healthy recipes.
Service lines	Specialties and service lines that the hospital offers; includes inpatient and ambulatory services.
Access	Features of the hospital that make receiving care easier for patients. This can be in the form of some convenience (location or hours of operation), insurances accepted, lists of physicians affiliated with the hospital, ways to interact with the hospital through social media, and online features, such as emergency department wait times and online bill pay.
Technology	Medical technology or equipment the hospital uses. Examples include robotic surgery and 3D mammography.
Research	Any form of inquiry into outcomes of disease. Includes formal clinical trials under review at the hospital and any highlighted research of individual staff members.
Patient education	Any information for patients on diseases, such as prevention, treatment, or disease management; includes classes, videos, and risk assessment guides in addition to literature.
Affiliations	Formal relationships with universities or hospital systems. This only includes hospitals that advertise an affiliation with a university or hospital system on their website.
Employment opportunities	Information on available jobs at the hospital, including job listings and online job applications.

Results

We estimated the percent of hospitals in the United States that use various categories of advertising on their website. We present whether the hospital had a focus on the category (indicated by having either a major focus or a minor focus) and whether there was any content relating to the category (major focus, minor focus, link, and search results, if applicable). Our findings are presented in [Table 3](#).

The major finding was that very few hospitals focused on price or quality information. Only 1.3% (6/474) of hospitals had a price focus on their website, 5.7% (27/574) had a quality focus, and 3.2% (15/474) had a patient safety focus. Instead, hospital website advertising was geared more toward service lines (56.5%, 268/474), access (49.6%, 235/474), awards (34.0%,

161/474), and amenities (30.8%, 146/474). The most common form of advertising that hospitals had was information on employment opportunities (92.6%, 439/474), which is not directly applicable to patients.

Having some information on price, however, did not mean that patients had full access to price. Of the 6 hospitals that had any price information in their focus area, only 1 had it as a major focus, although that major focus was limited to the price of 1 procedure (a \$99 calcium score screening). Of the 5 hospitals with a minor focus on price, 3 had information on cost estimates or ranges of common prices for common procedures, 1 had the price of 2 weight loss procedures, and 1 had the price of 1 heart scan.

Quality outcomes information was accessible on 40.9% (194/474) of websites. The information, however, was generally minimal and was often difficult to find or interpret on the website. Only 3 hospitals used quality data as their major focus, with 2 referencing Medicare's Hospital Compare data and the third, a cancer specialty hospital, showed cancer survival rates. Hospitals that included quality data often cited Hospital Compare data, but provided no comparison to other local hospitals. Others only provided data on a subset of outcomes measures, such as a general hospital only providing cardiovascular outcomes measures. Further, only 23.0% (109/474) of all hospitals had quality outcomes measurements on their site, with the remainder linking to external sources, primarily Hospital Compare or state quality reporting sites [13]. Often, data on quality were available if a consumer was willing to look for it, but hospitals were not actively competing on it, with 5.7% (27/474) having it as a major or minor focus.

Patient safety information was much less common than quality information, with only 20.3% (96/474) of hospitals having any such information available. The most common data was Hospital Compare data or Hospital Safety Score information from the

Leapfrog Group [14]. In all, 44.8% (43/96) of hospitals with patient safety information linked to offsite data. Patient safety was mentioned qualitatively at least as often as quality, but quantitative data were much less common; we suspect this is because no hospital wants to show data indicating they are unsafe for any of their patients.

The most common patient-directed advertisement method was to emphasize specific service lines the hospital offers; 56.5% (268/474) of hospitals focused on 1 or more service lines and 93.5% (443/474) had some information on service lines on their home page. A list of the specific service lines that hospitals had a major focus on is available in Table 4. The most common service lines to receive a major focus included specialties related to heart disease, cancer, women's services, and orthopedic surgery, which are often considered to be profit centers for hospitals [15]. Access (49.6%, 235/474) and amenities (30.8%, 146/474) were also very common focus categories, whereas customer satisfaction was not (2.7%, 13/474). These categories were focused around how the experience of care would be for the patients, as opposed to the actual reported experience of care.

Table 3. Percent of US hospitals that use various types of website advertising (N=474).

Type of advertising	Focus		Any information	
	n (%)	SE %	n (%)	SE %
Price	6 (1.3)	0.5	90 (19.0)	1.7
Quality				
Quality outcomes	27 (5.7)	1.0	194 (40.9)	2.1
Patient safety	15 (3.2)	0.8	96 (20.3)	1.8
Customer satisfaction	13 (2.7)	0.7	83 (17.5)	1.7
Awards and accreditation	161 (34.0)	2.1	222 (46.8)	2.2
Other factors				
Stories	83 (17.5)	1.7	117 (24.7)	1.9
Amenities	146 (30.8)	2.0	309 (65.2)	2.1
Services	268 (56.5)	2.2	443 (93.5)	1.1
Access	235 (49.6)	2.2	419 (88.4)	1.4
Technology	112 (23.6)	1.9	140 (29.5)	2.0
Research	16 (3.4)	0.8	62 (13.1)	1.5
Patient education	157 (33.1)	2.1	356 (75.1)	1.9
Any university affiliation			25 (5.3)	1.0
Any health system affiliation			233 (49.2)	2.2
Any employment information			439 (92.6)	1.1

Table 4. Hospital websites with a major focus on service lines.

Major service line	n	% of total hospitals (N=474)	% of hospitals with any major focus service line (n=170)
Hospitals with any major focus service line	170	35.9	100
Specific service lines			
Bariatric/weight loss/eating disorders	21	4.4	12.4
Cardiovascular disease	72	15.2	42.4
Emergency medicine	21	4.4	12.4
Labor and delivery	9	1.9	5.3
Multiple services	11	2.3	6.5
Neurology/neurosurgery	15	3.2	8.8
Obstetrics/gynecology	39	8.2	22.9
Oncology/cancer care	44	9.3	25.9
Orthopedic surgery	33	7.0	19.4
Other service line	54	31.8	11.4
Pediatrics	12	2.5	7.1
Radiology	13	2.7	7.6
Surgery	21	4.4	12.4
Wound care	9	1.9	5.3

Discussion

Overview

With so few hospitals focusing on price and quality information, it is apparent that hospitals are not actively competing for individual patients based on these factors. For a value-driven, consumer-based health system to hope to function, information on price and quality must be available. In our present system, however, there is insufficient information available for a consumer that wants to be engaged to adequately compare their options, rendering a consumer-driven system in its current form unviable.

Price Information

Nearly 19.0% (90/474) of hospitals do have some price information available on their website, but the information that is available is limited. Ohio, which had the most hospitals with some price information (86.4%, 19/22), has a law that requires hospitals to post price information on their websites, but this is limited to daily room charges and hospital charges for the 30 most common services in a variety of departments [3]. There are multiple challenges, however, with this law. First, the statute does not require hospitals to include information on physician services, supplies, or other nonhospital charges the patient may be billed for during their hospitalization. Second, it is limited to the specific hospital's most common procedures; therefore, the procedure a patient may be interested in may not be included on a hospital's charge list. Third, the service descriptions use technical language and there is no requirement to define the terms; for example, a patient comparing emergency department charges would have no direction how to estimate what level (1 through 5) their issue may be classified as or whether their need may require more or less than 31 to 74 minutes of critical care.

Fourth, and very significantly, there is no requirement to differentiate between charges and reimbursement rates. Hospitals charge rates have grown extensively over the past decades and significantly differ from actual payments that payers make for services rendered [16,17]. With no breakdown of what a patient may expect to pay, particularly in relation to their insurance, there is no way for consumers to appropriately compare costs.

Consumers are unable to adequately compare prices when fewer than 1 in 5 hospitals has any information on prices available. Of the hospital websites that do contain information on price, the information is limited to charges. No hospital provided information by insurer on how much a patient's out-of-pocket or total costs may actually be. Calculating out-of-pocket estimates may be difficult, but all hospitals have ready access to their chargemaster (a record of prices for all billable services) and, for most insurers, hospitals have ready access to the negotiated reimbursement rate for each service. At a minimum, hospitals could provide information on how much Medicare copays may be because that represents a large portion of their patient population with known copays. The information is available; it is just not being shared.

If this information is readily available to hospitals, why is it not available to patients? There are many potential reasons that a hospital may not share price information, including already having a ready supply of patients and because patients are not actively seeking this information. Although some hospitals are in competitive markets, many are in areas where they are the only hospital within comfortable travel distance, effectively giving them a regional monopoly on hospital services. Because rural hospitals are in less-populated areas, it is expected that they would also be in less-competitive markets as fewer hospitals will be close together. If this is the case, we would

expect urban hospitals to be more likely to have price information than rural hospitals. Although we found this to be the case (urban 22.0% vs rural 14.7%, $P=.03$), even urban hospitals appear to have little need to share price information. There may not be pressure on hospitals to provide this information because most patients may not care due to their health insurance shielding them from the effects of the price variation [18]. Patients may also rely on physicians to make hospital choices for them.

There are also reasons that may actually discourage hospitals from sharing information, including not wanting to hurt their bargaining position with insurers and because it limits their ability to price discriminate among individual patients. It is known that there are wide discrepancies between and within markets as to what hospitals are reimbursed for similar services [19-21]. If actual prices paid by different insurers were available, hospitals' negotiating power with those insurers would likely be weakened and their ability to price discriminate while negotiating with insurers or with individual purchasers would be diminished [22,23].

Quality

Similar to price, hospitals are not directly competing on quality outcomes. Although information on general outcomes is readily available through Hospital Compare, hospitals are not actively advertising their quality outcomes data to their patients. Hospitals attempt to convey quality via proxy measures or may feel no need to compete for patients based on quality.

A proxy representation of quality is anything that will imply a high level of care and good outcomes. This can be done in multiple ways, including advertising specific service lines, referencing external reviews of the facility, and by advertising technological advances. The most common approach is to focus on a specialty and add qualitative descriptions of how high quality outcomes and patient satisfaction are achieved.

A second proxy for quality is awards. These represent external recognition of the hospital, usually involving a specific service line, and assumes the external reviewer, because of an ability to evaluate the hospital's performance in a way that average consumers cannot, is in a position to make an objective pronouncement on the hospital's quality. Awards, however, have been criticized for not correlating well with objective outcomes and for methodological problems, such as being biased toward reputation [24,25]. Hospitals often focus on awards (34.0%, 161/474) and nearly half (46.8%, 222/474) mention them. Table 5 contains a breakdown of common awards. Other awards commonly cited included specialty society accreditations, local business awards (such as "Best Places to Work"), "Most Beautiful Hospital" awards, and others. Hospitals that listed any awards mentioned 4.6 different awards, on average.

Another proxy is use of technology. Although some technological advancements do improve care, others have not been shown to lead to better clinical outcomes while costing more [26,27]. Whether technological innovations always justify the costs is debated, but technology's ability to attract patients is well known [28,29].

Another possibility for the dearth of quality outcome focus is that hospitals do not feel that the information is a priority for most patients. If they felt that some patients were interested and they wanted to compete for these patients, then the information would be made available, but it would not be a focus. If this were the case, hospitals that are in areas that are more competitive would be more likely to have some quality information than those in less-competitive markets. Indeed, urban hospitals are much more likely to provide any quality information than rural hospitals (urban: 50.2%, 138/277; rural: 27.9%, 55/197, $P<.001$).

Table 5. Awards listed on hospital websites.

Award	% of total hospitals (N=474)	% of hospitals with any awards (n=222)
Any award	46.8	100.0
US News	8.6	18.5
Magnet	8.0	17.1
Leapfrog Group	3.2	6.8
Joint Commission	16.0	34.2
Thomson Reuters	6.1	13.1
Other awards	33.5	71.6

What Are Hospitals Competing On?

We grouped hospitals into 2 categories: those that do have an advertising focus (excluding affiliation or employment opportunities) and those that do not. In all, 84.0% (398/474) of hospitals do have some advertising focus on their website. Hospitals without a focus may not be competing for individuals at all. With narrowing networks, patients will tend to go to hospitals where their insurance is accepted, meaning the responsibility to evaluate costs and quality is relegated to the

insurer [22,30]. Although insurers are undoubtedly interested in quality outcomes at the population level when negotiating with providers, their primary interest is minimizing population costs, given their purchasing power [31]. This, however, negates the potential impact of individual consumer-directed care as the responsibility is moved to third-party insurers.

Of the hospitals that do have some advertising focus, 89.6% (354/398) focus on service lines, access, or amenities. A focus on services (67.1%, 267/398 of hospitals with any focus) implies

that the service line, among the others at the hospital, is exceptional. Without associated data on why it is exceptional, such as quality or cost information, the hospital is not competing for the value of the service, but the brand of the service. A focus on access or amenities (72.1%, 287/398) speaks to the experience of care, such as ease of receiving services or quality of facilities. The experience of care is 1 of the triple aims mentioned by the Institute for Healthcare Improvement, but this experience of care is not generally associated with better outcomes or lower costs [32,33].

This focus on patient experience, however, does make sense because that represents what many consumers are primarily interested in. A survey of commercially insured patients found the most important factor in choosing a hospital is patient experience, which was more than twice as important as clinical reputation [34]. Another study found that an increase in amenities, such as good food and attentive staff, lead to a significantly greater demand for the hospital among patients [35,36]. In recent years, there has also been unprecedented growth in patient experience-focused hospitals [37]. It is unclear whether increasing access to price and quality information will lead to significant changes in consumer preferences at the point of service as other factors are likely more important, such as established physician relationships, location, and the amenities of the hospital where the patient will stay. Immediately before a hospitalization, particularly when a patient is either in an emergency or suffering from the effects of a chronic condition, is not an ideal time to require patients to actively compare quality and price values between hospitals.

A health care system that increases value may not be achievable via a consumer-directed design at the point of service if consumers are more interested in factors not related to price or clinical outcomes immediately before choosing a hospital. A better approach to increase value is to redirect the consumer-based designs away from the point of service of care and instead incent consumers to purchase insurance based on lower prices and higher quality. This can be accomplished by increasing price and quality information relating to the networks that insurers have negotiated with. If patients can be steered toward lower-cost, higher-quality providers before they are ill

and are generally satisfied with their care, then they are likely to continue with that provider [38]. Focusing consumer-based reforms on the point of service may be too late to achieve meaningful improvements in clinical outcomes and decreases in health care costs.

Limitations

A limitation of this study is that it is focused on information available on hospital websites. There remains the possibility that hospitals have price or quality data accessible to patients at another source that is convenient to use, but is not mentioned on their websites. We feel that this is unlikely, but it is a possibility. Given the high proportion of consumers that use the Internet for health comparisons, we feel this limitation is minor. It is more likely, however, that hospitals advertise apart from their website (print, television, billboards, etc). Our findings are thus limited to advertisements on the hospitals' own websites.

Conclusions

For a consumer-driven health care system to lower health care costs and better health outcomes, information on price and quality must be available, consumers must choose providers based on those factors, and then providers must compete to improve on price and quality. There currently is not adequate information available for consumers to compare prices. Further, the minimal price information that exists is insufficient for a consumer to estimate how much their care may actually cost them out-of-pocket. There is more information available to consumers on quality measures from third parties, but hospitals are not actively competing on clinical outcomes of care. Hospitals do, however, compete on proxies for quality, including awards and by advertising medical technology, but these proxies do not always correlate with improved clinical outcomes.

Rather than prices or quality, hospitals are primarily competing on patient experience factors, such as amenities and conveniences. This may be because consumers are more interested in the experience of care at the time they are sick. A better approach is to encourage patients to choose low-cost, high-quality providers much earlier, such as when they purchase health insurance, rather than waiting until they are sick.

Conflicts of Interest

None declared.

References

1. Robinson JC. Health savings accounts--the ownership society in health care. *N Engl J Med* 2005 Sep 22;353(12):1199-1202. [doi: [10.1056/NEJMp058097](https://doi.org/10.1056/NEJMp058097)] [Medline: [16177245](https://pubmed.ncbi.nlm.nih.gov/16177245/)]
2. Feldman R, Parente ST, Abraham J, Christianson JB, Taylor R. Health savings accounts: early estimates of national take-up. *Health Aff (Millwood)* 2005;24(6):1582-1591 [FREE Full text] [doi: [10.1377/hlthaff.24.6.1582](https://doi.org/10.1377/hlthaff.24.6.1582)] [Medline: [16284032](https://pubmed.ncbi.nlm.nih.gov/16284032/)]
3. LAW Writer Ohio Laws and Rules. 2006. 3727.42 Price information list URL: <http://codes.ohio.gov/orc/3727.42> [accessed 2013-03-12] [WebCite Cache ID 6F43ryDGA]
4. State of Wisconsin. 2009 Wisconsin Act 146. 2010 Mar 23. URL: <http://docs.legis.wisconsin.gov/2009/related/acts/146.pdf> [accessed 2013-03-12] [WebCite Cache ID 6F43xFx7Y]
5. Porter ME. What is value in health care? *N Engl J Med* 2010 Dec 23;363(26):2477-2481. [doi: [10.1056/NEJMp1011024](https://doi.org/10.1056/NEJMp1011024)] [Medline: [21142528](https://pubmed.ncbi.nlm.nih.gov/21142528/)]

6. Gaynor M. What Do We Know About Competition and Quality in Health Care Markets?. Cambridge, MA: National Bureau of Economic Research; 2006 Jun. URL: <http://www.nber.org/papers/w12301.pdf> [accessed 2013-08-13] [[WebCite Cache ID 6IqICn9hc](#)]
7. kaiseredu.org. US Health Care Costs URL: <http://www.kaiseredu.org/issue-modules/us-health-care-costs/background-brief.aspx> [accessed 2013-03-12] [[WebCite Cache ID 6F449423t](#)]
8. Zickuhr K, Smith A. Digital differences. Washington, DC: Pew Internet & American Life Project; 2012 Apr 13. URL: http://alexa.pewinternet.com/~media/Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed 2013-03-12] [[WebCite Cache ID 6F44BJVWD](#)]
9. Understanding the role of the Internet in the lives of consumers: 2012 Digital Influence Index Annual Global Study. St Louis, MO: Fleishman Hilliard; 2012. URL: <http://push.fleishmanhillard.netdna-cdn.com/dii/2012-DII-White-Paper.pdf> [accessed 2013-03-12] [[WebCite Cache ID 6F44D7mRB](#)]
10. 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/)]
11. Data.Medicare.gov. Hospital General Information URL: <https://data.medicare.gov/dataset/Hospital-General-Information/v287-28n3> [accessed 2013-03-12] [[WebCite Cache ID 6F44F1Vav](#)]
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)]
13. Medicare.gov. Hospital compare URL: [http://www.medicare.gov/hospitalcompare/\(X\(1\)S\(uqwmkpv4hdbln4qrlkqjdbm\)\)/search.aspx?AspxAutoDetectCookieSupport=1](http://www.medicare.gov/hospitalcompare/(X(1)S(uqwmkpv4hdbln4qrlkqjdbm))/search.aspx?AspxAutoDetectCookieSupport=1) [accessed 2013-03-12] [[WebCite Cache ID 6F44Hlp05](#)]
14. Hospital Safety Score. How safe is your local hospital? URL: <http://www.hospitalsafetyscore.org/> [accessed 2013-03-12] [[WebCite Cache ID 6F44KADBb](#)]
15. Berenson RA, Bodenheimer T, Pham HH. Specialty-service lines: salvos in the new medical arms race. *Health Aff (Millwood)* 2006;25(5):w337-w343 [[FREE Full text](#)] [doi: [10.1377/hlthaff.25.w337](https://doi.org/10.1377/hlthaff.25.w337)] [Medline: [16868008](https://pubmed.ncbi.nlm.nih.gov/16868008/)]
16. Tompkins CP, Altman SH, Eilat E. The precarious pricing system for hospital services. *Health Aff (Millwood)* 2006;25(1):45-56 [[FREE Full text](#)] [doi: [10.1377/hlthaff.25.1.45](https://doi.org/10.1377/hlthaff.25.1.45)] [Medline: [16403744](https://pubmed.ncbi.nlm.nih.gov/16403744/)]
17. Melnick GA, Fonkyk K. Hospital pricing and the uninsured: do the uninsured pay higher prices? *Health Aff (Millwood)* 2008 Mar;27(2):w116-w122 [[FREE Full text](#)] [doi: [10.1377/hlthaff.27.2.w116](https://doi.org/10.1377/hlthaff.27.2.w116)] [Medline: [18252736](https://pubmed.ncbi.nlm.nih.gov/18252736/)]
18. Harrington C, Estes C. Health Policy: Crisis and Reform in the US Health Care Delivery System. Sudbury, MA: Jones and Bartlett; 2008.
19. Melnick GA, Zwanziger J, Bamezai A, Pattison R. The effects of market structure and bargaining position on hospital prices. *Journal of Health Economics* 1992 Oct;11(3):217-233. [doi: [10.1016/0167-6296\(92\)90001-H](https://doi.org/10.1016/0167-6296(92)90001-H)]
20. Ginsburg PB. Research brief: Wide variation in hospital and physician payment rates evidence of provider market power. Washington, DC: Center for Studying Health System Change; 2010 Nov. URL: <http://www.hschange.com/CONTENT/1162/1162.pdf> [accessed 2013-03-12] [[WebCite Cache ID 6F44R03z7](#)]
21. Robinson JC. Hospital market concentration, pricing, and profitability in orthopedic surgery and interventional cardiology. *Am J Manag Care* 2011 Jun;17(6 Spec No):e241-e248 [[FREE Full text](#)] [Medline: [21756018](https://pubmed.ncbi.nlm.nih.gov/21756018/)]
22. Reinhardt UE. The pricing of U.S. hospital services: chaos behind a veil of secrecy. *Health Aff (Millwood)* 2006 Jan;25(1):57-69 [[FREE Full text](#)] [doi: [10.1377/hlthaff.25.1.57](https://doi.org/10.1377/hlthaff.25.1.57)] [Medline: [16403745](https://pubmed.ncbi.nlm.nih.gov/16403745/)]
23. Pauly MV, Burns LR. Price transparency for medical devices. *Health Aff (Millwood)* 2008;27(6):1544-1553 [[FREE Full text](#)] [doi: [10.1377/hlthaff.27.6.1544](https://doi.org/10.1377/hlthaff.27.6.1544)] [Medline: [18997210](https://pubmed.ncbi.nlm.nih.gov/18997210/)]
24. Halasyamani LK, Davis MM. Conflicting measures of hospital quality: ratings from "Hospital Compare" versus "Best Hospitals". *J Hosp Med* 2007 May;2(3):128-134. [doi: [10.1002/jhm.176](https://doi.org/10.1002/jhm.176)] [Medline: [17549759](https://pubmed.ncbi.nlm.nih.gov/17549759/)]
25. Sehgal AR. The role of reputation in U.S. News & World Report's rankings of the top 50 American hospitals. *Ann Intern Med* 2010 Apr 20;152(8):521-525. [doi: [10.7326/0003-4819-152-8-201004200-00009](https://doi.org/10.7326/0003-4819-152-8-201004200-00009)] [Medline: [20404383](https://pubmed.ncbi.nlm.nih.gov/20404383/)]
26. Kang CM, Kim DH, Lee WJ, Chi HS. Conventional laparoscopic and robot-assisted spleen-preserving pancreatectomy: does da Vinci have clinical advantages? *Surg Endosc* 2011 Jun;25(6):2004-2009. [doi: [10.1007/s00464-010-1504-1](https://doi.org/10.1007/s00464-010-1504-1)] [Medline: [21136089](https://pubmed.ncbi.nlm.nih.gov/21136089/)]
27. Pugh TJ, Choi S, Nguyen QN, Gillin MT, Ron Zhu X, Palmer MB, et al. Proton beam therapy for the treatment of prostate cancer. *Practical Radiation Oncology* 2013 Apr;3(2):e87-e94. [doi: [10.1016/j.pro.2012.05.010](https://doi.org/10.1016/j.pro.2012.05.010)]
28. Pauly MV. Competition and new technology. *Health Aff (Millwood)* 2005;24(6):1523-1535 [[FREE Full text](#)] [doi: [10.1377/hlthaff.24.6.1523](https://doi.org/10.1377/hlthaff.24.6.1523)] [Medline: [16284024](https://pubmed.ncbi.nlm.nih.gov/16284024/)]
29. Devers KJ, Brewster LR, Casalino LP. Changes in hospital competitive strategy: a new medical arms race? *Health Serv Res* 2003 Feb;38(1 Pt 2):447-469 [[FREE Full text](#)] [Medline: [12650375](https://pubmed.ncbi.nlm.nih.gov/12650375/)]
30. Kaiser Family Foundation, Health Research & Educational Trust. Employer Health Benefits: 2012 Annual Survey. 2012. URL: <http://ibpis.com/blog/wp-content/uploads/2012/05/2011-KFF.pdf> [accessed 2013-03-12] [[WebCite Cache ID 6F44YS0VS](#)]
31. Reinhardt UE. New York Times: Economix. 2010 Nov 26. How medicare sets hospital prices: a primer URL: <http://economix.blogs.nytimes.com/2010/11/26/how-medicare-sets-hospital-prices-a-primer/> [accessed 2013-03-12] [[WebCite Cache ID 6F44g9bYt](#)]

32. Lee DS, Tu JV, Chong A, Alter DA. Patient satisfaction and its relationship with quality and outcomes of care after acute myocardial infarction. *Circulation* 2008 Nov 4;118(19):1938-1945 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.108.792713](https://doi.org/10.1161/CIRCULATIONAHA.108.792713)] [Medline: [18936325](https://pubmed.ncbi.nlm.nih.gov/18936325/)]
33. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med* 2012 Mar 12;172(5):405-411. [doi: [10.1001/archinternmed.2011.1662](https://doi.org/10.1001/archinternmed.2011.1662)] [Medline: [22331982](https://pubmed.ncbi.nlm.nih.gov/22331982/)]
34. Grote KD, Newman JRS, Sutaria SS. McKinsey&Company. 2007 Nov. A better hospital experience: hospitals must learn what commercially insured patients and their physicians look for when choosing facilities—and how to deliver it URL: http://www.mckinseyquarterly.com/A_better_hospital_experience_2081 [accessed 2013-03-12] [WebCite Cache ID [6F44inown](https://www.webcitation.org/6F44inown)]
35. Goldman D, Romley JA. Hospitals as hotels: the role of patient amenities in hospital demand. Cambridge, MA: National Bureau of Economic Research; 2008 Dec. URL: <http://www.nber.org/papers/w14619.pdf> [accessed 2013-08-13] [WebCite Cache ID [6IqI6Cj7T](https://www.webcitation.org/6IqI6Cj7T)]
36. Goldman DP, Vaiana M, Romley JA. The emerging importance of patient amenities in hospital care. *N Engl J Med* 2010 Dec 2;363(23):2185-2187. [doi: [10.1056/NEJMp1009501](https://doi.org/10.1056/NEJMp1009501)] [Medline: [21121830](https://pubmed.ncbi.nlm.nih.gov/21121830/)]
37. Conte A, Fábregas L. TribLIVE.com. 2011 Dec 18. Medical building boom increasing Americans' health care costs URL: <http://triblive.com/investigative/specialprojects/codegreen/1983696-84/hospital-health-hospitals-care-upmc-beds-construction-facilities-patients-bed> [accessed 2013-03-12] [WebCite Cache ID [6F44qJVNT](https://www.webcitation.org/6F44qJVNT)]
38. Marquis MS, Davies AR, Ware JE. Patient satisfaction and change in medical care provider: a longitudinal study. *Med Care* 1983 Aug;21(8):821-829. [Medline: [6888031](https://pubmed.ncbi.nlm.nih.gov/6888031/)]

Abbreviations

CBSA: core-based statistical area

CMS: Centers for Medicare and Medicaid Services

HCAHPS: Hospital Consumer Assessment of Healthcare Providers and Systems

HCRIS: Healthcare Cost Report Information System

SE: standard error

VA: Veterans Affairs

Edited by G Eysenbach; submitted 09.04.13; peer-reviewed by D Jaffe, M Green; comments to author 15.06.13; revised version received 19.06.13; accepted 20.06.13; published 29.08.13.

Please cite as:

Muhlestein DB, Wilks CEA, Richter JP

Limited Use of Price and Quality Advertising Among American Hospitals

J Med Internet Res 2013;15(8):e185

URL: <http://www.jmir.org/2013/8/e185/>

doi: [10.2196/jmir.2660](https://doi.org/10.2196/jmir.2660)

PMID: [23988296](https://pubmed.ncbi.nlm.nih.gov/23988296/)

©David B Muhlestein, Chrisanne E A Wilks, Jason P Richter. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 29.08.2013. 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

The MOonitoring Resynchronization dEVICES and CARdiac patiENTS (MORE-CARE) Randomized Controlled Trial: Phase 1 Results on Dynamics of Early Intervention With Remote Monitoring

Giuseppe Boriani¹, MD, PhD; Antoine Da Costa², MD; Renato Pietro Ricci³, MD; Aurelio Quesada⁴, MD, PhD; Stefano Favale⁵, MD; Saverio Iacopino⁶, MD; Francesco Romeo⁷, MD; Arnaldo Risi⁸, PhD; Lorenza Mangoni di Stefano⁸, MS; Xavier Navarro⁹, MD; Mauro Biffi¹, MD; Massimo Santini³, MD; Haran Burri¹⁰, MD; On Behalf Of The MORE-CARE Investigators¹¹

¹Institute of Cardiology, Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, S Orsola-Malpighi University Hospital, Bologna, Italy

²University Hospital, St Etienne, France

³San Filippo Neri Hospital, Rome, Italy

⁴University General Hospital, Valencia, Spain

⁵University Hospital, Bari, Italy

⁶Anthea Hospital, Bari, Italy

⁷Fondazione Policlinico Tor Vergata Hospital, Rome, Italy

⁸Medtronic Clinical Research Institute, EMEA Regional Clinical Center, Rome, Italy

⁹Medtronic Ibérica, Scientific & Clinical Department, Madrid, Spain

¹⁰Cardiology Service, University Hospital of Geneva, Geneva, Switzerland

¹¹see Multimedia Appendix 3 for collaborators

Corresponding Author:

Giuseppe Boriani, MD, PhD

Institute of Cardiology

Department of Experimental, Diagnostic and Specialty Medicine

University of Bologna, S Orsola-Malpighi University Hospital

Via Massarenti, 9

Bologna,

Italy

Phone: 39 051 349858

Fax: 39 051 344859

Email: giuseppe.boriani@unibo.it

Abstract

Background: Remote monitoring (RM) in patients with advanced heart failure and cardiac resynchronization therapy defibrillators (CRT-D) may reduce delays in clinical decisions by transmitting automatic alerts. However, this strategy has never been tested specifically in this patient population, with alerts for lung fluid overload, and in a European setting.

Objective: The main objective of Phase 1 (presented here) is to evaluate if RM strategy is able to reduce time from device-detected events to clinical decisions.

Methods: In this multicenter randomized controlled trial, patients with moderate to severe heart failure implanted with CRT-D devices were randomized to a Remote group (with remote follow-up and wireless automatic alerts) or to a Control group (with standard follow-up without alerts). The primary endpoint of Phase 1 was the delay between an alert event and clinical decisions related to the event in the first 154 enrolled patients followed for 1 year.

Results: The median delay from device-detected events to clinical decisions was considerably shorter in the Remote group compared to the Control group: 2 (25th-75th percentile, 1-4) days vs 29 (25th-75th percentile, 3-51) days respectively, $P=.004$. In-hospital visits were reduced in the Remote group (2.0 visits/patient/year vs 3.2 visits/patient/year in the Control group, 37.5% relative reduction, $P<.001$). Automatic alerts were successfully transmitted in 93% of events occurring outside the hospital in the Remote group. The annual rate of all-cause hospitalizations per patient did not differ between the two groups ($P=.65$).

Conclusions: RM in CRT-D patients with advanced heart failure allows physicians to promptly react to clinically relevant automatic alerts and significantly reduces the burden of in-hospital visits.

Trial Registration: Clinicaltrials.gov NCT00885677; <http://clinicaltrials.gov/show/NCT00885677> (Archived by WebCite at <http://www.webcitation.org/6IkCJ7NF>).

(*J Med Internet Res* 2013;15(8):e167) doi:[10.2196/jmir.2608](https://doi.org/10.2196/jmir.2608)

KEYWORDS

cardiac resynchronization therapy; heart failure; alerts; remote monitoring; telemedicine

Introduction

Modern cardiac resynchronization therapy defibrillators (CRT-D) are equipped with reliable diagnostics able to provide a series of alerts, including lung fluid accumulation [1], occurrence of atrial fibrillation (AF) [2], or technical issues. Early diagnosis and intervention may play a crucial role in minimizing major cardiovascular events and reducing hospitalization. Several major device manufacturers offer remote monitoring (RM) capabilities [3,4] with the aim of reducing regular follow-up visits [5] and unnecessary interim visits, or of dealing with the more complex perspective of disease management [6]. RM allows physicians to remotely access patient data and to be notified of clinical events by means of the automatic transmission of alert messages. Previous trials such as TRUST [7] and CONNECT [8] have shown that RM is safe and reduces delay in detection of events such as AF. However, these trials either excluded [7] or included only a minority of patients with biventricular defibrillators (CRT-D) [8] and did not include an alert on lung fluid accumulation, which is potentially useful in the context of heart failure management. In addition, the aforementioned trials [7,8] as well as the EVOLVO trial [9] were not strictly focused on NYHA III-IV heart failure patients, a setting where reduction of morbidity and access to hospitals may have a great significance for both the patient and the health care system.

MORE-CARE is a multicenter randomized trial conducted in Europe and designed in two phases [10]. Phase 1 was aimed at evaluating whether RM of CRT-D patients could shorten the time from onset of a clinically relevant event to clinical decisions in comparison with standard management (scheduled in-office visits only). The second phase of MORE-CARE is currently ongoing and is targeted at assessing whether clinical decision making guided by RM exerts a positive impact on patient outcome (death from any cause, cardiovascular and device-related hospitalization) in comparison to standard care [10]. This paper presents the results of Phase 1 of the study.

Methods

Remote Monitoring With CareLink Network

CareLink Network, as detailed in [Multimedia Appendix 1](#), is a platform for remote monitoring of implantable cardiac devices, which consists of implantable devices provided with wireless telemetry, CareLink monitor (CLM), and the CareLink website (CW). The system allows patients to send comprehensive implant device data to their clinic from home or any location where CareLink is available and that is equipped with an analog

telephone line or cellular connection. The organization of the platform and how patients and health care professionals interact are shown in [Figure 1](#). More specifically, every communication between implantable cardiac devices and CLM is based on the Conexus wireless telemetry technology, which uses the Medical Implant Communications Service (MICS) radio frequency band (between 402 and 405 MHz), specifically designed for medical devices and targeted to reduce the risk of interfering with other users of the same band. In case of either scheduled or device-detected event transmission, device information are collected by CLM using the aforementioned wireless communication system and then transmitted via a private data network by means of a phone line connection. Unique credentials created at the time of manufacture and stored in each monitor are used to authenticate the monitor to the CareLink network. Health care providers can analyze the transmitted data via the Internet by using a Web browser to access the CW. The site is also used to enroll clinic users and patients and perform other administrative duties. The Medtronic CareLink network operates on the Windows operating system with database support based on Microsoft's SQL Server software.

With regard to data protection, health care professionals access patient data via the Internet through a connection to a Microsoft Internet Information Server (IIS). In addition, every user session is protected by means of the 128-bit Secure Sockets Layer (SSL) encryption system.

Study Design

The MORE-CARE study is an international, prospective, multicenter randomized controlled trial in patients with a Medtronic CRT-D, designed to compare disease management guided by RM with the CareLink network with standard clinical practice. The trial design has been reported in detail elsewhere [10], and the flow chart is shown in [Figure 2](#). In summary, patients in sinus rhythm with de novo implantation of CRT-D for systolic heart failure with NYHA class III/IV (and a left ventricular ejection fraction (LVEF) <35% were randomized 1:1 to wireless RM Remote group or to a Control group. Patients in the Remote group had in-office visits at baseline and at 8 months, and remote follow-ups performed at 4 and 12 months with activation of automatic alerts (for AF, lung fluid via OptiVol monitoring, device integrity, ineffective shocks, or inactivated ventricular fibrillation (VF) detection/therapy). Control group patients had in-office visits performed at baseline and at every 4 months. Audible alerts for device integrity issues or for inactivated VF detection/therapy were activated in both groups. The institutional ethics committees approved the protocol at all 32 centers involved. All patients were enrolled after signing an informed consent form.

Figure 1. Remote monitoring system platform and interactions between health care professionals and patient.

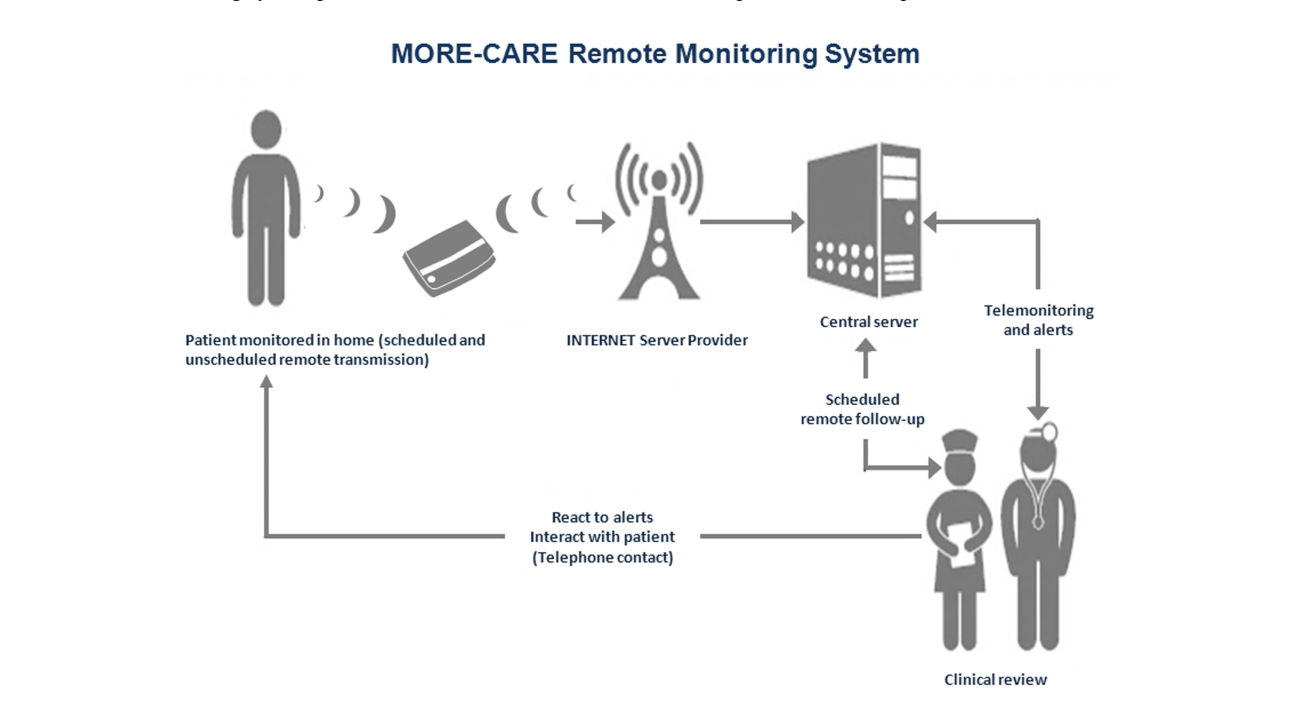
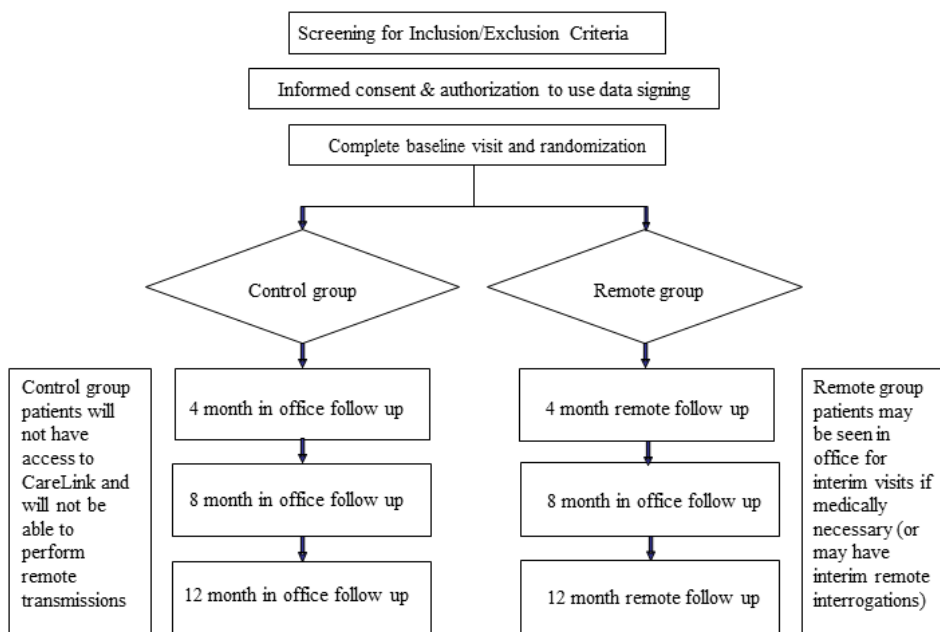


Figure 2. Scheduling of follow-up in the Remote group (with RM and in-office follow-up) and in the Control group (with in-office follow-up only).



Programming of Diagnostic Features

In the Remote group, automatic RM alerts for lung fluid accumulation (using OptiVol fluid status monitoring), atrial tachyarrhythmia (AT/AF), lead and device integrity, or inactivated VF detection/therapy were turned on at baseline using standardized predefined thresholds, which could later be modified at the physicians' discretion. All audible alerts were disabled with the exception of those related to lead and device integrity and programming.

Due to technological aspects of the RM system, any device-detected event for AT/AF, fast ventricular rate during AT/AF and shock triggers a remote notification of the episode and automatically disables the corresponding alert, which may be re-enabled only by means of an in-office device check (alert re-arming). For patients in the Control group, only audible alerts for system integrity and programming issues were enabled. Tachycardia detection and treatment were performed according to the standard practice of the individual centers, as well as optimization of atrio-ventricular (AV) and ventricular-ventricular (VV) intervals for CRT (see [Multimedia Appendix 2](#)).

Objectives

The primary endpoint of Phase 1 (whose participants are reported in (see [Multimedia Appendix 3](#)) was to demonstrate a reduction of the time from onset of an actionable device-detected event to a clinical decision. A clinical decision was defined by at least one of the following: change in drug therapy, device reprogramming, patient education (specific advice on salt, fluid check, exercise, behavior, etc), as well as planning of hospital admission for other interventions (eg, electrical cardioversion, radiofrequency ablation, etc). For appropriate clinical decisions and patient management, physicians were aided by specific flowcharts suggesting clinical actions for each type of alert (system performance alerts, device shocks, AF "rhythm control" strategy, AF "rate control" strategy, and OptiVol events). The flowcharts are shown in [Multimedia Appendices 4-8](#). The time of a device-detected event was defined as the day on which the criterion for an alert was satisfied (for subjects in the Control group, this was taken into account as well, even though the alert was switched off). If no actions were taken following the acknowledgement of a device event, it was not considered for primary endpoint adjudication.

The primary endpoint was defined as the delay between onset of the actionable device-detected event to a clinical decision related to that event. Endpoints were validated by a blinded Endpoint Adjudication Committee (see [Multimedia Appendix 3](#)). Phase 1 included as secondary objectives an exploratory analysis of the impact of RM on quality of life (by means of the Minnesota Living with Heart Failure Questionnaire) and clinical status (measured by the Clinical Composite Score).

Data Collection

Phase 1 of the MORE-CARE study required the collection of clinical and device data at scheduled and unscheduled visits (either ambulatory or remotely performed). The onset day of all device-detected events in the Control group patients was determined by evaluating the device memory information

collected at each in-office visit. For the Remote group patients, dates of alert transmissions and dates when alerts were reviewed by either nurses or physicians were collected by each study center; the date of alert transmission was considered as the onset day of a device-detected event. In addition, all clinical actions by physicians on the basis of either device memory information or device alert notification were collected and dated; the clinical actions date to treat a specific event was considered as the end date of such event and used to determine the delay between the event onset to a clinical decision. Therefore, for each event, 3 different times have been collected: (1) time of event detection (start time), (2) time to when the physician was able to review the event (time of event revision), and (3) time to clinical actions taken to treat the event (time to clinical decision).

Statistical Analysis

In accordance with the sample size estimation reported in the study design [10], the sample size requirement was 77 subjects per study arm. In the current report of Phase 1 results, data analysis was performed according to the intention-to-treat principle. The analysis for the primary endpoint includes all randomized subjects who experienced at least one event. Similar to a previous study [8], if a patient experienced multiple events of a specific type (eg, ≥ 6 h AT/AF burdens in a day) between 2 consecutive evaluations, only the first of these was paired with the next device interrogation/visit and counted toward the analysis. For every patient, an average time from event onset to clinical decision was calculated for each type of event and entered in the analysis per event type.

Descriptive statistics are reported as mean \pm 1 standard deviation for normally distributed continuous variables, or medians with 25th-75th percentiles for skewed variables. Normality of distribution was tested by the Kolmogorov-Smirnov test. Absolute and relative frequencies are reported for categorical variables. Continuous Gaussian variables were compared by the Student's *t* test for independent samples, while skewed distributions were compared using the Mann-Whitney nonparametric test. Differences in proportions were compared by applying Chi-square analysis. Rates of events were computed per 100 person years, as number of occurred events out of patient exposure time and reported separately for each arm. The exposure time was computed from the date of randomization to the date of the last available information for each patient, either dropped out or died. Rates were compared by means of the Comparison Incidence Rates (Large Sample) Test. An alpha-level of .05 was considered for each test. All statistical analyses were performed by using SAS 9.3 version software.

Results

Study Population

A total of 154 patients were enrolled from May 2009 through April 2010 from 32 centers in 6 different countries (France, Hungary, Israel, Italy, Spain, and Switzerland). The final patient cohort object of analysis comprised 148 patients (76 in the Remote group and 72 in the Control group see [Figure 3](#)). Demographic data and clinical parameters of the population under analysis were similar between the study arms at the time

of enrollment (Table 1). The median follow-up duration was 12.0 months (25th-75th percentile, 11.8-12.6 months) with 1709 cumulative months of follow-up.

Device-Detected Events and Remote Transmission

Of the 148 patients, 105 (71%) experienced at least 1 event satisfying the criteria for triggering a device alert: 57 patients (75%) in the Remote group and 48 patients (67%) in the Control group ($P=.28$). There were 166 alerts in the Remote group and 114 episodes matching the alert criteria in the Control group. OptiVol threshold crossing and AT/AF burden were the most frequent events (Table 2). In particular, the observed rate of OptiVol events was similar between the 2 arms (1.6 events/year in the Remote group, 1.5 events/year in the Control group, $P=.59$), while the rate of events of AT/AF burden and fast ventricular rate during AF episodes was higher in the Remote group (0.7 events/year) compared to the Control arm (0.2 events/year, $P<.001$).

Of the 166 alerts in the Remote group, 144 (87%) were successfully transmitted. The remaining 22 alerts (13%) in the Remote group were not successfully transmitted because the patient was admitted to hospital before transmission in 11 cases, the monitor was not set up in 8 cases, there were connection problems with the phone line in 2 cases, and the patient was not at home in 1 case. Discounting the alerts with failed transmission due to hospital admission, successful transmission occurred in 144/155 (93%) of events.

For the 166 alerts in the Remote group, the median delay between triggering of the alert to when the event was reviewed (remotely or in-office) was 3 days (25th-75th percentile, 1-10 days) compared with the median time of 37 days (25th-75th percentile, 14-71 days) for the 114 device events of the Control group ($P<.001$, Table 2).

Time From Actionable Device-Detected Event to Clinical Decisions

Overall, 56 device-detected events led to at least one clinical decision taken by a physician participating in the study protocol (37 events in 23 Remote group patients and 19 events in 15 Control group patients). Device-detected episodes of lung fluid accumulation (53%, $n=30$) and AT/AF burden above the pre-specified limit (34%, $n=19$) were the most frequent events leading to consequent clinical actions (Table 3). The median time from the event onset to related clinical decisions was significantly shorter in the Remote group vs the Control group (2 days vs 29 days, $P=.004$, 93% relative reduction, Figure 4).

Of all 56 clinical decisions consequent to device-detected events, 44 of them involved 1 clinical action, 11 had 2 clinical actions and only 1 case had 3 clinical actions performed at the same time (Table 4). Therefore, a total of 69 clinical actions (43 in

the Remote group and 26 in the Control group) were taken on the basis of the above-mentioned 56 device-detected events. In both groups, the majority of clinical actions were medication changes (58% in the Remote group and 50% in the Control group, $P=.78$). Device re-programming constituted 20% of clinical actions in the Remote group and 23% in the Control group ($P=.65$), while hospitalizations were decided only for 6 patients in the Control group and none of the patients in the Remote group ($P=.001$, Figure 5).

In-hospital Visits

There were a total of 375 scheduled follow-ups: 189 for the 76 Remote group patients (125 remote follow-ups and 64 in-office visits) and 186 in-office visits for the 72 patients of the Control group. Overall, taking into account both scheduled and unscheduled visits (in a referral clinic) plus emergency department visits (with or without subsequent hospitalization) a 37.5% reduced burden was observed in the Remote group (144 visits, corresponding to 2.00 visits/year vs 225 visits, corresponding to 3.20 visits/year in the Control group, $P<.001$, Figure 6).

Hospital Admissions

During the follow-up of Phase 1, there were 19 hospitalizations for various causes (related to 18 patients) in the Remote group and 22 hospitalizations (related to 16 patients) in the Control group ($P=.65$).

Quality of Life and Clinical Status

The patient's quality of life was assessed by means of the Minnesota Living with Heart Failure Questionnaire. Baseline values were comparable between the Remote group (41; 25th-75th percentile, 16-62) and the Control group (40; 25th-75th percentile, 18-53, $P=.38$). The change in quality of life from baseline to the 8th month was not different for the Remote group (-17; 25th-75th percentile, -32 to -2) compared to the Control group (-10; 25th-75th percentile, -23 to 0, $P=.45$). The change in clinical status during the trial from enrollment to the 12-month follow-up was similar in both groups according to the Clinical Composite Score. In the Remote group, 54% of patients were defined as "improved", 35% as "unchanged", and 11% as "worsened", while in the Control group these values were 48%, 38%, and 14% respectively ($P=.69$).

Deaths and Study Exits

During the course of Phase 1, 7 patients died for the following reasons: heart failure (3 patients in the Remote group and 1 patient in the Control group), complications after aortic surgery (1 patient in the Remote group), stroke (1 patient in the Control group), and chronic kidney disease (1 patient in the Remote group). Furthermore, 9 patients exited prematurely from the trial for reasons reported in Figure 3.

Table 1. Demographics and baseline clinical parameters.

Subject characteristics	Control group, n=72	Remote group, n=76	P value
Male gender, n (%)	54 (75.0)	55 (72.4)	.72
Age, years	68±9	67±10	.63
Ischemic heart disease, n (%)	32 (44.4)	39 (51.3)	.40
NYHA functional classification at implant, n (%)			.59
Class III	70 (98.5)	70 (94.5)	
Class IV	1 (1.4)	4 (5.4)	
Previous myocardial infarction, n (%)	30 (41.7)	39 (51.3)	.24
Hypertension, n (%)	27 (37.5)	30 (39.5)	.81
History of coronary artery intervention, n (%)	21 (29.2)	22 (29.0)	.98
History of valvular surgery, n (%)	8 (11.1)	5 (6.6)	.33
Chronic Obstructive Pulmonary Disease (COPD), n (%)	15 (20.8)	13 (17.0)	.56
AF paroxysmal, n (%)	8 (11.1)	15 (19.7)	.15
AF persistent, n (%)	1 (1.4)	4 (5.3)	.37
Implant for primary prevention, n (%)	65 (90.3)	68 (89.5)	.87
Diabetes, n (%)	23 (32.4)	26 (35.6)	.68
QRS (ms)	148±30	155±25	.11
LVEF (%)	27±7	27±6	.69
ACE inhibitor or ARB, n (%)	60 (83.3)	64 (84.2)	.89
β-blocker, n (%)	63 (87.5)	66 (86.8)	.91
Diuretic, n (%)	69 (95.8)	71 (93.4)	.52
Antiarrhythmic agents, n (%)	17 (23.6)	18 (23.7)	.99

Figure 3. Phase 1 follow-up experience flow-chart.

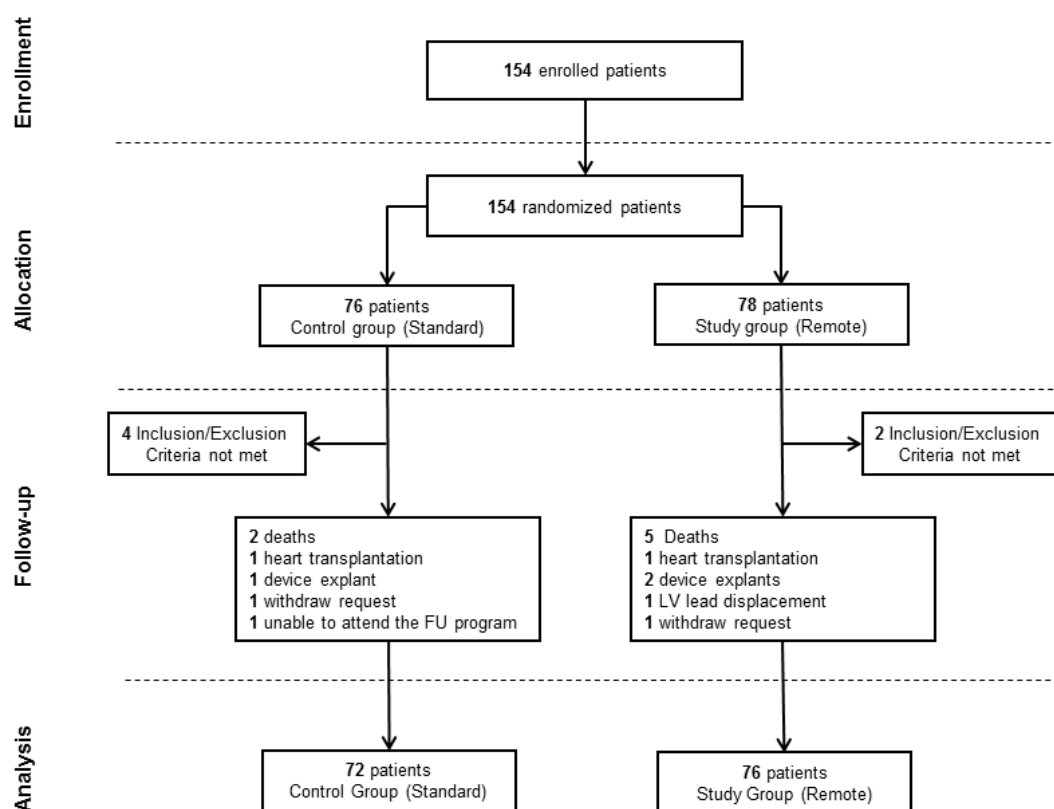


Table 2. Device-detected events—delays from a device-detected event to review of the alert (including alerts with unsuccessful transmissions in the Remote group, which were evaluated during an in-office visit).

Device-detected event	Total # of device-detected events (# of patients)		# of days from device-detected event to event reviewing Median (25 th -75 th percentile)		P value
	Control group	Remote group	Control group	Remote group	
Lead impedances out of range	5 (4)	2(1)	6 (0-22)	12 (0-23)	1.000
VF detection/therapy off	0 (0)	3 (2)	-	0 (0-0)	-
AT/AF burden: at least 6 hours of AT/AF in a single day	9 (7)	39 (12)	51 (5-59)	2 (1-7)	.002
Fast V rate during AT/AF: Mean V rate of at least 100 bpm a day with at least 6 hrs of AT/AF	5 (4)	7 (3)	57 (2-68)	1 (0-6)	.19
Number of shocks delivered in an episode (at least two)	1 (1)	5 (5)	0 (0-0)	0 (0-1)	.73
OptiVol threshold crossing for lung fluid accumulation	94 (47)	110 (55)	39 (22-72)	4 (1-12)	<.001
Total	114 (48)	166 (57)	37 (14-71)	3 (1-10)	<.001

Table 3. Delay in clinical decisions—delays from a device-detected event to a clinical decision.

Device-detected event	# of actionable device-detected events followed by a clinical decision (# of patients)		# of days from actionable device-detected event to clinical decision Median (25 th -75 th percentile)		P value
	Control group	Remote group	Control group	Remote group	
Lead impedances out of range	2 (2)	0 (0)	24 (13-35)	-	-
VF detection/therapy off	0 (0)	0 (0)	-	-	-
AT/AF burden: at least 6 hrs of AT/AF in a single day	5 (5)	14 (7)	51 (0-59)	2 (1-3)	.51
Fast V rate during AT/AF: Mean V rate of at least 100 bpm a day with at least 6 hrs of AT/AF	0 (0)	2 (2)	-	9 (1-17)	-
Number of shocks delivered in an episode (at least two)	1 (1)	2 (2)	0 (0-0)	0 (0-0)	.50
OptiVol threshold crossing for lung fluid accumulation	11 (9)	19 (15)	29 (9-31)	3 (1-6)	.002
Total	19 (15)	37 (23)	29 (3-51)	2 (1-4)	.004

Table 4. Distribution of clinical actions among all actionable device-detected events (19 in the Control group and 37 in the Remote group) followed by a clinical decision.

Actionable device-detected event	Patient education		Device programming		Medication changes		Hospitalizations		Laboratory tests	
	Cont. group	Rem. group	Cont. group	Rem. group	Cont. group	Rem. group	Cont. group	Rem. group	Cont. group	Rem. group
AT/AF burden: at least 6 hrs of AT/AF in a single day (Control group n=5, Remote group n=14)	6	4	2	2	4	8	2	0	0	1
Fast V rate during AT/AF: Mean V rate of at least 100 bpm a day with at least 6 hrs of AT/AF (Control group n=0, Remote group n=2)	0	0	0	0	0	1	0	0	0	1
# of shocks delivered in an episode (at least 2) (Control group n=1, Remote group n=2)	0	0	0	2	0	1	1	0	0	0
OptiVol threshold crossing for lung fluid accumulation (Control group n=11, Remote group n=19)	1	6	3	4	9	13	2	0	0	0
Lead impedances out of range (Control group n=2, Remote group n=0)	0	0	1	0	0	0	1	0	0	0

Figure 4. Time from device events to clinical decisions for the phase 1 primary end-points (19 in Control group and 37 in Remote group); box-and-whisker plots show the quartiles with the medians labeled, and the whiskers extended to the lower and the upper adjacent value; plus symbols show the outside values.

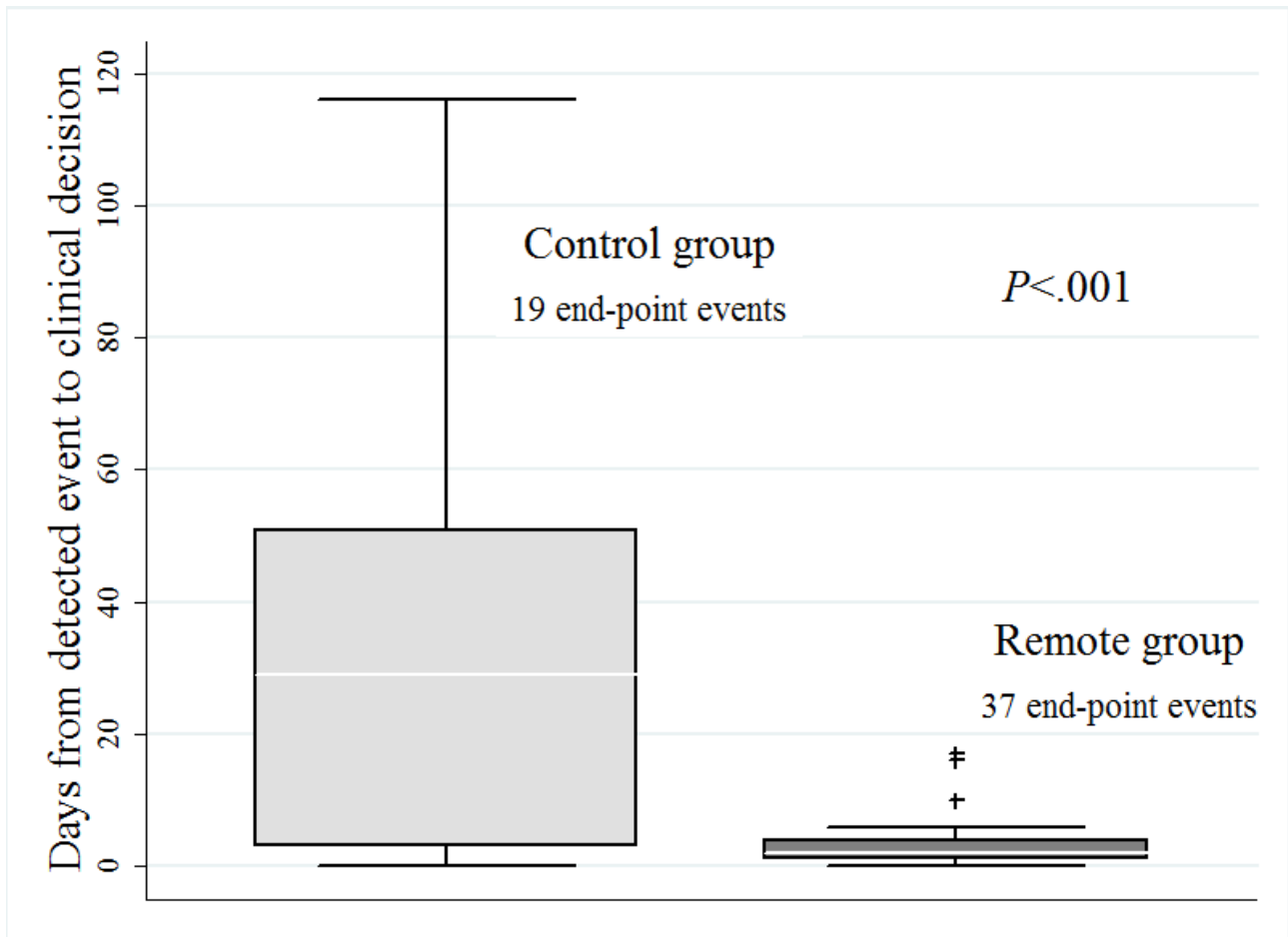


Figure 5. Distribution of specific clinical actions related to device-detected events in the Remote group (n=43) and in the Control group (n=26) respectively.

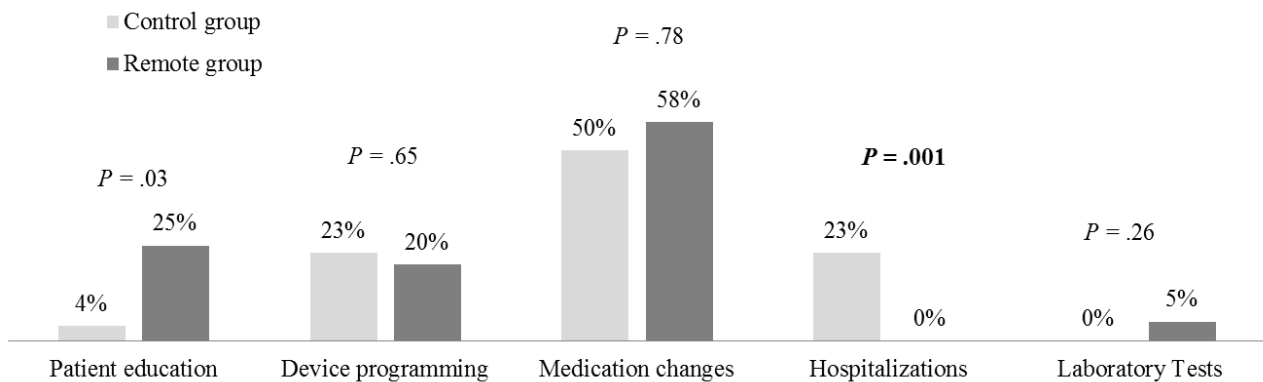
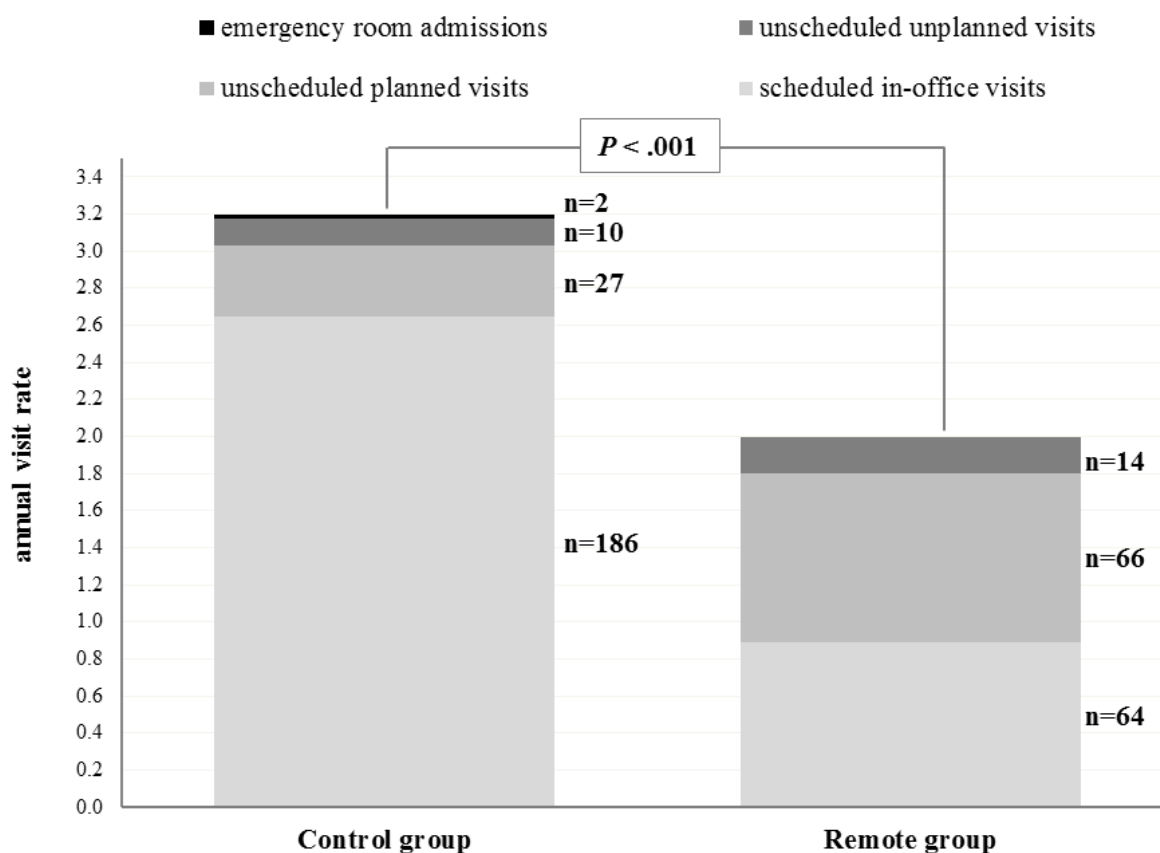


Figure 6. Annual rates per randomization group of scheduled visits (in-office visits performed as per protocol requirement), unscheduled planned (in-office visits not required by the protocol, not patient initiated), and unplanned visits (in-office visits not required by the protocol, patient initiated) and emergency room admissions (for each randomization group and for each type of visit, the total number of occurrences is displayed beside the corresponding bar).



Discussion

Key Findings

The MORE-CARE trial is the first European, large-scale randomized study evaluating disease management guided by RM (including lung fluid overload alerts) in a population comprised exclusively of CRT-D patients with advanced heart failure. The main finding of Phase 1 of the study that we report here is that a wireless RM strategy permits physicians to take clinical decisions 27 days sooner as compared to standard in-office care, while reducing the total number of in-hospital visits.

Comparison With Other Trials

Several recent studies [7-9,11-14] have demonstrated the benefits associated with remote patient monitoring in terms of early detection of relevant events as well as reduction of in-office follow-up visits. However, none of these studies was dedicated specifically to CRT-D patients with advanced heart failure, and most were conducted in the United States. Recently, the results of the EVOLVO trial were reported, showing a significant reduction in emergency visits in patients on RM [9]. The EVOLVO trial, however, differed significantly from MORE-CARE in that the patient population was enrolled

exclusively in one region of Italy, with a mixture of ICD and CRT-D patients of whom >80% had NYHA class I/II heart failure, with activation of audible alerts in the control arm, and without evaluation of delay in medical decisions [9,15].

The Phase 1 results of MORE-CARE revealed a median 27-days reduction in delay from actionable event detection to medical decision for the Remote group compared to the Control group. This reduction was even greater than the median 17-days reduction observed in the CONNECT trial, most probably due to different in-office follow-up intervals (3 months vs 4 months respectively). The delay in reviewing events was considerably shorter in the Remote arm. For AF burden for example, this delay was reduced by a median of 49 days, which compared favorably to the median of 34 days observed in the TRUST study (probably also in part due to differences in follow-up intervals). Delay from the AF alert to clinical action was thereby significantly reduced (2 vs 51 days). This is particularly important in the case of new AF episodes because timely diagnosis and prompt treatment may minimize thromboembolic complications [16-19] and may prevent heart failure (HF) decompensation and inappropriate shocks. In the EVOLVO study, the overall reduction in delay to reviewing of alerts was 23 days, compared to 34 days in MORE-CARE [10]. The difference may be explained by activated audible alerts in the

EVOLVO control group. Audible alerts may indeed be useful for device integrity issues but may lead to increased rates of hospital visits or admissions for OptiVol alerts, as shown in the DOT-HF study [15,20]. These data underline the importance of RM with automatic wireless alerts notified to physicians, rather than monitoring based on audible alerts delivered to patients, as a strategy for avoiding unnecessary hospital visits and use of resources.

In addition to timely reviewing of alerts and clinical decision making, the prerequisite for improved patient outcome resulting from RM is the reliable transmission of these alerts. The CONNECT trial showed that 45% of all alerts were not transmitted, mainly due to the monitor not being set up [8]. The percentage of unsuccessful alert transmissions in MORE-CARE was considerably less, 13% in total or only 7% if failed transmission due to hospital admission is discounted. Patient education and logistics for implementing RM may have partly accounted for these differences.

In our study, the number of detected AF events was higher in the Remote group compared to the Control group. This can be explained by taking into account the higher percentage of paroxysmal AF at baseline in the former group compared to the latter and by the fact that alerts are re-armed by in-office device interrogations. Therefore, a single episode of persistent AF may generate multiple alerts if the patient undergoes several in-office device interrogations that are prompted by the alerts.

Monitoring of HF decompensation with OptiVol has been evaluated in several trials [20-23], none of which involved RM. The CONNECT trial did not include the use of the OptiVol algorithm as the feature is not available as an alert in the United States. As reported by a previous study [1], intrathoracic impedance appears to be inversely correlated with pulmonary capillary pressure and its decrease may serve as an early notification to identify patients at high risk of impending exacerbation of congestive heart failure [21,22]. The PARTNERS-HF study [23] has shown that integrating different diagnostics (eg, presence of arrhythmias, patient activity, heart rate variability, and nocturnal heart rate) significantly improved the ability to identify patients at risk of heart failure events, beyond the use of intrathoracic impedance alone. In addition to reliable alert transmission and prompt response to these alerts, appropriate interpretation of diagnostics is likely to affect the ability of an RM strategy to improve patient outcome. We included specific clinical pathways in the MORE CARE trial

to favor appropriate alert management. Following analysis of device diagnostics and patient phone contact (specified in the clinical pathway), only a minority (17%) of the OptiVol alerts resulted in clinical action.

Our results showed a significant reduction in the rate of in-office visits in the Remote group compared to the Control group, despite an increased number of unscheduled visits resulting from alerts, which is in line with previous trials that did not specifically target patients with CRT-D [7-9]. This suggests a potential benefit of RM in terms of health care logistics and costs in this patient population with advanced heart failure. However, a more thorough analysis, taking into account all costs of in-office and remote activities, is needed to confirm this aspect.

Even though not significant, there was a trend in greater improvement in quality of life at 8 months in the Remote group compared to the Control group. It is likely that the effect of CRT (rather than follow-up strategy) may be preponderant in improving quality of life early after CRT implantation. Other endpoints such as hospitalizations and mortality were not significantly different between the groups.

Limitations

Phase 1 of MORE-CARE was not powered for evaluating the impact of RM on cardiovascular and device-related hospitalizations and mortality, which are being studied in the second phase of the trial. There were only a few cases of system integrity alerts because of the limited 1-year follow-up. These aspects will be better evaluated with the longer observation period of the ongoing trial.

Conclusion

The Phase 1 results of the MORE-CARE randomized study indicate that RM allows a significant reduction in delay from event onset to clinical decisions. In spite of a reduced number of in-hospital visits compared to patients with standard follow-up, we found no significant differences among the groups in terms of quality of life and clinical status. The impact of RM on clinical aspects of disease management in heart failure patients needs to be assessed in the second phase of the trial. Finally, these findings are reported for the first time in a European setting, in a study cohort consisting entirely of CRT-D patients with advanced heart failure in whom remote disease management included alerts for lung fluid overload.

Acknowledgments

The authors thank Mr. Francesco De Seta, Mr. Andrea Grammatico, and Ms. Alessandra Denaro for their assistance in the preparation of this manuscript, and Marco Vimercati for his support in the data management. The MORE-CARE study is funded by Medtronic, Inc (Minneapolis, MN).

Conflicts of Interest

Arnaldo Risi, Lorenza Mangoni di S Stefano, and Xavier Navarro are employees of affiliates of Medtronic Inc. Prof. Boriani has received speaker fees from Medtronic (small amount); Dr. Ricci has received minor consultancy fees from Medtronic and Biotronik; Prof. Santini has received research grants from Medtronic, and St. Jude. Dr. Burri has received grants for research and

fellowships support and speaker fees from Biotronik, Boston Scientific, Medtronic, St-Jude Medical, and Sorin. The other authors report no conflict.

Multimedia Appendix 1

Carelink system description and general workflow for Carelink scheduled and alert transmission.

[[PDF File \(Adobe PDF File\), 276KB - jmir_v15i8e167_app1.pdf](#)]

Multimedia Appendix 2

Implantable cardiac device concepts.

[[PDF File \(Adobe PDF File\), 50KB - jmir_v15i8e167_app2.pdf](#)]

Multimedia Appendix 3

List of Steering Committee members, Endpoint Adjudication Committee members, investigators (and centers) participating in the MORE-CARE Phase 1.

[[PDF File \(Adobe PDF File\), 77KB - jmir_v15i8e167_app3.pdf](#)]

Multimedia Appendix 4

Flowchart for managing system performance alerts.

[[PNG File, 15KB - jmir_v15i8e167_app4.png](#)]

Multimedia Appendix 5

Flowchart for atrial fibrillation episodes. “Rhythm control” strategy.

[[PNG File, 21KB - jmir_v15i8e167_app5.png](#)]

Multimedia Appendix 6

Flowchart for atrial fibrillation episodes. “Rate control” strategy.

[[PNG File, 19KB - jmir_v15i8e167_app6.png](#)]

Multimedia Appendix 7

Flowchart for managing the possible fluid accumulation alerts (Optivol).

[[PNG File, 25KB - jmir_v15i8e167_app7.png](#)]

Multimedia Appendix 8

Flowchart for managing device shock and ICD therapies alerts.

[[PNG File, 21KB - jmir_v15i8e167_app8.png](#)]

Multimedia Appendix 9

CONSORT-EHEALTH checklist V1.6.2 [24].

[[PDF File \(Adobe PDF File\), 989KB - jmir_v15i8e167_app9.pdf](#)]

References

1. Yu CM, Wang L, Chau E, Chan RH, Kong SL, Tang MO, et al. Intrathoracic impedance monitoring in patients with heart failure: correlation with fluid status and feasibility of early warning preceding hospitalization. *Circulation* 2005 Aug 9;112(6):841-848 [[FREE Full text](#)] [doi: [10.1161/CIRCULATIONAHA.104.492207](https://doi.org/10.1161/CIRCULATIONAHA.104.492207)] [Medline: [16061743](https://pubmed.ncbi.nlm.nih.gov/16061743/)]
2. Lewalter T, Morgan J, Halimi F, Lip G, Dagues N, Blomström-Lundqvist C, Scientific Initiative Committee, European Heart Rhythm Association. Monitoring in the management of atrial fibrillation. *Europace* 2012 Apr;14(4):591-592 [[FREE Full text](#)] [doi: [10.1093/europace/eus073](https://doi.org/10.1093/europace/eus073)] [Medline: [22454456](https://pubmed.ncbi.nlm.nih.gov/22454456/)]

3. Schoenfeld MH, Compton SJ, Mead RH, Weiss DN, Sherfese L, Englund J, et al. Remote monitoring of implantable cardioverter defibrillators: a prospective analysis. *Pacing Clin Electrophysiol* 2004 Jun;27(6 Pt 1):757-763. [doi: [10.1111/j.1540-8159.2004.00524.x](https://doi.org/10.1111/j.1540-8159.2004.00524.x)] [Medline: [15189530](#)]
4. Lazarus A. Remote, wireless, ambulatory monitoring of implantable pacemakers, cardioverter defibrillators, and cardiac resynchronization therapy systems: analysis of a worldwide database. *Pacing Clin Electrophysiol* 2007 Jan;30 Suppl 1:S2-S12. [doi: [10.1111/j.1540-8159.2007.00595.x](https://doi.org/10.1111/j.1540-8159.2007.00595.x)] [Medline: [17302706](#)]
5. Boriani G, Auricchio A, Klersy C, Kirchhof P, Brugada J, Morgan J, European Heart Rhythm Association, Eucomed. Healthcare personnel resource burden related to in-clinic follow-up of cardiovascular implantable electronic devices: a European Heart Rhythm Association and Eucomed joint survey. *Europace* 2011 Aug;13(8):1166-1173 [FREE Full text] [doi: [10.1093/europace/eur026](https://doi.org/10.1093/europace/eur026)] [Medline: [21345922](#)]
6. Boriani G, Diemberger I, Martignani C, Biffi M, Valzania C, Bertini M, et al. Telecardiology and remote monitoring of implanted electrical devices: the potential for fresh clinical care perspectives. *J Gen Intern Med* 2008 Jan;23 Suppl 1:73-77 [FREE Full text] [doi: [10.1007/s11606-007-0355-5](https://doi.org/10.1007/s11606-007-0355-5)] [Medline: [18095049](#)]
7. Varma N, Epstein AE, Irimpen A, Schweikert R, Love C, TRUST Investigators. Efficacy and safety of automatic remote monitoring for implantable cardioverter-defibrillator follow-up: the Lumos-T Safely Reduces Routine Office Device Follow-up (TRUST) trial. *Circulation* 2010 Jul 27;122(4):325-332 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.110.937409](https://doi.org/10.1161/CIRCULATIONAHA.110.937409)] [Medline: [20625110](#)]
8. Crossley G, Boyle A, Vitense H, Sherfese L, Mead RH. Trial design of the clinical evaluation of remote notification to reduce time to clinical decision: the Clinical evaluation Of remote NotificatioN to rEduCe Time to clinical decision (CONNECT) study. *Am Heart J* 2008 Nov;156(5):840-846. [doi: [10.1016/j.ahj.2008.06.028](https://doi.org/10.1016/j.ahj.2008.06.028)] [Medline: [19061696](#)]
9. Landolina M, Perego GB, Lunati M, Curnis A, Guenzati G, Vicentini A, et al. Remote monitoring reduces healthcare use and improves quality of care in heart failure patients with implantable defibrillators: the evolution of management strategies of heart failure patients with implantable defibrillators (EVOLVO) study. *Circulation* 2012 Jun 19;125(24):2985-2992 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.111.088971](https://doi.org/10.1161/CIRCULATIONAHA.111.088971)] [Medline: [22626743](#)]
10. Burri H, Quesada A, Ricci RP, Boriani G, Davinelli M, Favale S, et al. The MOnitoring Resynchronization dEVICES and CARdiac patiEnts (MORE-CARE) study: rationale and design. *Am Heart J* 2010 Jul;160(1):42-48. [doi: [10.1016/j.ahj.2010.04.005](https://doi.org/10.1016/j.ahj.2010.04.005)] [Medline: [20598971](#)]
11. Mabo P, Victor F, Bazin P, Ahres S, Babuty D, Da Costa A, COMPAS Trial Investigators. A randomized trial of long-term remote monitoring of pacemaker recipients (the COMPAS trial). *Eur Heart J* 2012 May;33(9):1105-1111 [FREE Full text] [doi: [10.1093/eurheartj/ehr419](https://doi.org/10.1093/eurheartj/ehr419)] [Medline: [22127418](#)]
12. Halimi F, Clémenty J, Attuel P, Dessenne X, Amara W, OEDIPE trial Investigators. Optimized post-operative surveillance of permanent pacemakers by home monitoring: the OEDIPE trial. *Europace* 2008 Dec;10(12):1392-1399 [FREE Full text] [doi: [10.1093/europace/eun250](https://doi.org/10.1093/europace/eun250)] [Medline: [18775878](#)]
13. Crossley GH, Chen J, Choucair W, Cohen TJ, Gohn DC, Johnson WB, PREFER Study Investigators. Clinical benefits of remote versus transtelephonic monitoring of implanted pacemakers. *J Am Coll Cardiol* 2009 Nov 24;54(22):2012-2019. [doi: [10.1016/j.jacc.2009.10.001](https://doi.org/10.1016/j.jacc.2009.10.001)] [Medline: [19926006](#)]
14. Elsner C, Sommer P, Piorkowski C, Taborsky M, Neuser H, Bytesnik J, et al. A Prospective Multicenter Comparison Trial of home monitoring against regular follow-up in MADIT II: additional visitscost impact. In A. Murray, editor. In: 33rd Conference on Computers in Cardiology. Washington, DC: IEEE; 2006:421-424.
15. van Veldhuisen DJ, Maass AH. Telemonitoring of outpatients with heart failure: a search for the holy grail? *Circulation* 2012 Jun 19;125(24):2965-2967 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.112.118141](https://doi.org/10.1161/CIRCULATIONAHA.112.118141)] [Medline: [22626742](#)]
16. Ricci RP, Morichelli L, Santini M. Remote control of implanted devices through Home Monitoring technology improves detection and clinical management of atrial fibrillation. *Europace* 2009 Jan;11(1):54-61 [FREE Full text] [doi: [10.1093/europace/eun303](https://doi.org/10.1093/europace/eun303)] [Medline: [19011260](#)]
17. Boriani G, Diemberger I, Mantovani V, Biffi M, Martignani C. Remote monitoring of patients with an implanted device and patients' outcomes: the potential for "win-win" dynamics. *J Cardiovasc Electrophysiol* 2009 Nov;20(11):1252-1254. [doi: [10.1111/j.1540-8167.2009.01576.x](https://doi.org/10.1111/j.1540-8167.2009.01576.x)] [Medline: [19682165](#)]
18. Boriani G, Gasparini M, Landolina M, Lunati M, Proclemer A, Lonardi G, ClinicalService cardiac centres. Incidence and clinical relevance of uncontrolled ventricular rate during atrial fibrillation in heart failure patients treated with cardiac resynchronization therapy. *Eur J Heart Fail* 2011 Aug;13(8):868-876. [doi: [10.1093/eurjhf/hfr046](https://doi.org/10.1093/eurjhf/hfr046)] [Medline: [21558331](#)]
19. Boriani G, Santini M, Lunati M, Gasparini M, Proclemer A, Landolina M, Italian ClinicalService Project. Improving thromboprophylaxis using atrial fibrillation diagnostic capabilities in implantable cardioverter-defibrillators: the multicentre Italian ANGELS of AF Project. *Circ Cardiovasc Qual Outcomes* 2012 Mar 1;5(2):182-188 [FREE Full text] [doi: [10.1161/CIRCOUTCOMES.111.964205](https://doi.org/10.1161/CIRCOUTCOMES.111.964205)] [Medline: [22373906](#)]
20. van Veldhuisen DJ, Braunschweig F, Conraads V, Ford I, Cowie MR, Jondeau G, DOT-HF Investigators. Intrathoracic impedance monitoring, audible patient alerts, and outcome in patients with heart failure. *Circulation* 2011 Oct 18;124(16):1719-1726 [FREE Full text] [doi: [10.1161/CIRCULATIONAHA.111.043042](https://doi.org/10.1161/CIRCULATIONAHA.111.043042)] [Medline: [21931078](#)]

21. Abraham WT, Compton S, Haas G, Foreman B, Canby RC, Fishel R, FAST Study Investigators. Intrathoracic impedance vs daily weight monitoring for predicting worsening heart failure events: results of the Fluid Accumulation Status Trial (FAST). *Congest Heart Fail* 2011 Apr;17(2):51-55. [doi: [10.1111/j.1751-7133.2011.00220.x](https://doi.org/10.1111/j.1751-7133.2011.00220.x)] [Medline: [21449992](https://pubmed.ncbi.nlm.nih.gov/21449992/)]
22. Small RS, Wickemeyer W, Germany R, Hoppe B, Andrulli J, Brady PA, et al. Changes in intrathoracic impedance are associated with subsequent risk of hospitalizations for acute decompensated heart failure: clinical utility of implanted device monitoring without a patient alert. *J Card Fail* 2009 Aug;15(6):475-481. [doi: [10.1016/j.cardfail.2009.01.012](https://doi.org/10.1016/j.cardfail.2009.01.012)] [Medline: [19643357](https://pubmed.ncbi.nlm.nih.gov/19643357/)]
23. Whellan DJ, Ousdigian KT, Al-Khatib SM, Pu W, Sarkar S, Porter CB, PARTNERS Study Investigators. Combined heart failure device diagnostics identify patients at higher risk of subsequent heart failure hospitalizations: results from PARTNERS HF (Program to Access and Review Trending Information and Evaluate Correlation to Symptoms in Patients With Heart Failure) study. *J Am Coll Cardiol* 2010 Apr 27;55(17):1803-1810. [doi: [10.1016/j.jacc.2009.11.089](https://doi.org/10.1016/j.jacc.2009.11.089)] [Medline: [20413029](https://pubmed.ncbi.nlm.nih.gov/20413029/)]
24. 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

ACE: angiotensin-converting enzyme
ARB: angiotensin receptor blocker
AF: atrial fibrillation
AT: atrial tachycardia
AV: atrio-ventricular
CLM: CareLink monitor
COPD: chronic obstructive pulmonary disease
CRT-D: cardiac resynchronization therapy defibrillators
CW: CareLink website
HF: heart failure
LVEF: left ventricular ejection fraction
MICS: Medical Implant Communications Service
NYHA: New York Heart Association
OptiVol: fluid status monitoring tool
QRS: combination of three of the graphical deflections seen on a typical electrocardiogram
RM: remote monitoring
VF: ventricular fibrillation
V rate: ventricular rate
VT: ventricular tachycardia
VV: ventricular-ventricular

Edited by G Eysenbach; submitted 21.03.13; peer-reviewed by D Smithard, D Hayn, G Schreier; comments to author 12.04.13; revised version received 20.05.13; accepted 09.06.13; published 21.08.13.

Please cite as:

Boriani G, Da Costa A, Ricci RP, Quesada A, Favale S, Iacopino S, Romeo F, Risi A, Mangoni di S Stefano L, Navarro X, Biffi M, Santini M, Burri H, On Behalf Of The MORE-CARE Investigators

The MOnitoring Resynchronization dEVICES and CARdiac patiEnts (MORE-CARE) Randomized Controlled Trial: Phase 1 Results on Dynamics of Early Intervention With Remote Monitoring

J Med Internet Res 2013;15(8):e167

URL: <http://www.jmir.org/2013/8/e167/>

doi: [10.2196/jmir.2608](https://doi.org/10.2196/jmir.2608)

PMID: [23965236](https://pubmed.ncbi.nlm.nih.gov/23965236/)

©Giuseppe Boriani, Antoine Da Costa, Renato Pietro Ricci, Aurelio Quesada, Stefano Favale, Saverio Iacopino, Francesco Romeo, Arnaldo Risi, Lorenza Mangoni di S Stefano, Xavier Navarro, Mauro Biffi, Massimo Santini, Haran Burri, On Behalf Of The MORE-CARE Investigators. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.08.2013. 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

Web-Based Versus Traditional Paper Questionnaires: A Mixed-Mode Survey With a Nordic Perspective

Lena Hohwü¹, MHS; Heidi Lyshol²; Mika Gissler^{3,4}, PhD; Stefan Hrafn Jonsson⁵, PhD; Max Petzold⁶, PhD; Carsten Obel¹, PhD

¹Department of Public Health, Section for General Practice, Aarhus University, Aarhus C, Denmark

²Department of Health Statistics, Norwegian Institute of Public Health, Oslo, Norway

³THL National Institute for Health and Welfare, Helsinki, Finland

⁴NHV Nordic School of Public Health, Gothenburg, Sweden

⁵University of Iceland and The Directorate of Health, Reykjavik, Iceland

⁶Sahlgrenska Academy, Centre for Applied Biostatistics, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:

Lena Hohwü, MHS

Department of Public Health

Section for General Practice

Aarhus University

Bartholins Allé 2

Aarhus C, 8000

Denmark

Phone: 45 87167942

Fax: 45 86124788

Email: lena.hohwu@alm.au.dk

Abstract

Background: Survey response rates have been declining over the past decade. The more widespread use of the Internet and Web-based technologies among potential health survey participants suggests that Web-based questionnaires may be an alternative to paper questionnaires in future epidemiological studies.

Objective: To compare response rates in a population of parents by using 4 different modes of data collection for a questionnaire survey of which 1 involved a nonmonetary incentive.

Methods: A random sample of 3148 parents of Danish children aged 2-17 years were invited to participate in the Danish part of the NordChild 2011 survey on their children's health and welfare. NordChild was conducted in 1984 and 1996 in collaboration with Finland, Iceland, Norway, and Sweden using mailed paper questionnaires only. In 2011, all countries used conventional paper versions only except Denmark where the parents were randomized into 4 groups: (1) 789 received a paper questionnaire only (paper), (2) 786 received the paper questionnaire and a log-in code to the Web-based questionnaire (paper/Web), (3) 787 received a log-in code to the Web-based questionnaire (Web), and (4) 786 received log-in details to the Web-based questionnaire and were given an incentive consisting of a chance to win a tablet computer (Web/tablet). In connection with the first reminder, the nonresponders in the paper, paper/Web, and Web groups were also present with the opportunity to win a tablet computer as a means of motivation. Descriptive analysis was performed using chi-square tests. Odds ratios were used to estimate differences in response rates between the 4 modes.

Results: In 2011, 1704 of 3148 (54.13%) respondents answered the Danish questionnaire. The highest response rate was with the paper mode (n=443, 56.2%). The other groups had similar response rates: paper/Web (n=422, 53.7%), Web (n=420, 53.4%), and Web/tablet (n=419, 53.3%) modes. Compared to the paper mode, the odds for response rate in the paper/Web decreased by 9% (OR 0.91, 95% CI 0.74-1.10) and by 11% (OR 0.89, 95% CI 0.73-1.09) in the Web and Web/tablet modes. The total number of responders for NordChild declined from 10,291 of 15,339 (67.09%) in 1984 and 10,667 of 15,254 (69.93%) in 1996 to 7805 of 15,945 (48.95%) in 2011 with similar declines in all 5 Nordic countries.

Conclusions: Web-based questionnaires could replace traditional paper questionnaires with minor effects on response rates and lower costs. The increasing effect on the response rate on participants replying for a nonmonetary incentive could only be estimated within the 2 Web-based questionnaire modes before the first reminder. Alternative platforms to reach higher participation rates

in population surveys should reflect the development of electronic devices and the ways in which the population primarily accesses the Internet.

(*J Med Internet Res* 2013;15(8):e173) doi:[10.2196/jmir.2595](https://doi.org/10.2196/jmir.2595)

KEYWORDS

mixed-mode survey; patient participation rate; Web-based; paper; questionnaires; nonmonetary incentive

Introduction

Background

Epidemiological studies have seen response rates decline by approximately 1% per year in many countries in recent years [1-3]. In general, a low response rate may increase the risk that respondents differ systematically from nonrespondents, in which case the results may not be representative of the study population [4]. This is not a serious problem in studies testing a specific causal hypothesis [3]. However, differential selection of respondents may seriously bias the results in descriptive and cross-sectional studies, and it is desirable to develop methods to maximize the response rate to achieve optimum external validity.

Response Rates in Paper- and Web-Based Questionnaires

The paper questionnaire has been the epidemiological mode of choice for collecting survey data so far, but with the increasing use of the Internet, Web-based questionnaires may be an obvious alternative. Web-based questionnaires have been shown to lower data collection costs [5,6], which is attractive especially in large population-based surveys.

The response rates in most studies so far, however, have been reported to be lower in Web-based questionnaires than in paper-based questionnaires [5,7], but the opposite has also been reported [8]. In a questionnaire survey on patients' experiences with breast cancer care, no significant difference was observed between the response rates of a mailed paper questionnaire only (64.0%) and an online questionnaire followed by a paper reminder (60.5%) [5]. A Danish questionnaire survey reported a statistically significantly higher total response rate in a paper-and-pencil group (76.5%) than in a group with access to the questionnaire via log-on to the Internet (64.2%) [7]. A study comparing mixed-mode (paper or online) and Web-based questionnaires exploring fertility issues among female childhood cancer survivors found a 6% higher participation rate in the Web-based mode (89%) than in the mixed mode (83%) [8].

Incentives

A number of previous studies suggest that monetary or lottery incentives increase response rates and that such incentives may be used to raise participant representativeness [4,9-12]. They showed an improvement in the response rate of 2 percentage points by the use of a US \$5 versus a US \$2 incentive, and that use of a £10 gift voucher gave 45% higher odds of responding than use of no incentive. The respondent in Web-based questionnaires may simply skip difficult items, which increases the risk of incompletely filled-in questionnaires, although this problem has also been noticed in paper questionnaires. The

inclination to skip items and the resulting lower rate of fully completed questionnaires has been seen predominantly in Web-based versions, but the problem has also been noticed in paper questionnaires [5,13]. However, an incentive both reduces item nonresponse and improves participation among participants with lower education levels [14].

Access to the Internet

The rapid growth in access to the Internet in developed countries has decreased the coverage differential between paper- and Web-based questionnaires and, thus, the risk of selection bias by using the Internet for questionnaire purposes [1]. In 2012, 99% of Danish couples with children had access to a computer and the Internet at home compared with 86% of all Danish families in general [15]. In Denmark, there is free public access to the Internet at all libraries, which ensures 100% access to the Internet for the whole population.

The questionnaire survey "Health and welfare among children and young people in the Nordic countries" (NordChild) was previously conducted in 1984 (NordChild1984) and 1996 (NordChild1996). The response rate increased from 67.09% in 1984 to 69.93% in 1996 [16]. Because of the general tendency observed in other studies [1,3], we expected the NordChild 2011 attrition to be considerably lower. In light of the advancement of Web-based technology and of its growing reliability, the Internet may be considered an obvious means for questionnaire data collection or at least a mode supplementary to the traditional paper questionnaire, not least because of its comparative advantages in terms of lower costs and simpler logistics [5,6]. Adding an incentive to the Internet-based questionnaire is thought to increase the response rate and to improve the quality of the answers [9]. To our knowledge, little is known about whether the presence of a single chance of receiving a nonmonetary incentive improves the response rate. The aim of the present paper is to compare response rates in a questionnaire survey by using 4 different modes of data collection, including 1 with a nonmonetary incentive.

Methods

The Danish Survey

A total of 3200 parents of children aged 2 to 17 years living in Denmark in 2011 were randomly selected to participate in the third NordChild questionnaire survey. There were 200 children in each of the 16 year groups between 2 and 17 years of age: 100 girls and 100 boys. The random selection of 1 child per family was managed by the Danish National Board of Health. All citizens in Denmark have a unique 10-digit personal identification number. The parents and their addresses were identified through the Danish Civil Registration System [17], and persons who had prohibited the use of their addresses for

marketing purposes or scientific studies were excluded before the selection [18]; in 2009, this group accounted for 12.8% of the Danish population [19]. The Danish questionnaire consisted of 73 questions with subquestions, and the paper version was 28 pages long. The study was approved by the Danish National Board of Health. For the analysis, the unique 10-digit personal identification number was encrypted.

A total of 52 invited respondents were excluded in the analysis; 5 announced that they did not wish to participate in the survey, 1 because of difficulties understanding written Danish, and 46 because the child had turned 18 years of age after the sample was selected and before the questionnaire was mailed.

The 3200 invited children were randomly allocated into 4 equal modes, and after exclusion they (N=3148) were allocated as following: (1) 789 received a paper questionnaire only (paper), (2) 786 received the paper questionnaire as well as a log-in code to the Web-based questionnaire (paper/Web), (3) 787 received log-in information to the Web-based questionnaire only (Web), and (4) 786 received log-in details to the Web-based questionnaire plus an incentive consisting of a chance to win a tablet computer (Web/tablet). Overall, the paper and paper/Web groups were categorized as paper versions of the questionnaire and the Web and Web/tablet as Web-based questionnaires. We choose to allocate the incentive in the Web-based questionnaires because the Web and Web/tablet modes are comparable. The Web-based questionnaire was a multipage design using SurveyXact [20] and had the same questions as the paper version did. The respondents of the Web-based questionnaire could answer the questions in several rounds, and submit it after the last question. A 12-character log-in code to the Web-based questionnaire had to be keyed in every time if it was not completed in a single round.

The data collection ran over a 4-month period, starting on June 6 and closing on October 6, 2011. All 3200 were invited by mail. Two reminders were mailed to all those who had not responded 4 and 12 weeks after the invitation to participate, on July 6 and August 31, respectively. Nonresponders at both the first and the second reminder were offered online participation only, but could receive a paper version of the questionnaire if they requested it. In connection with the first reminder, the nonresponders in the paper, paper/Web, and Web groups also were given the opportunity to win a tablet computer as means of motivation. In all, 2 tablet computers were distributed in the data collection period; the first to the responders in the Web/tablet mode, and the second to the responders who responded after the initial reminder in the paper, paper/Web, and Web modes.

Costs per responder were estimated. The estimation included distribution and collection of the questionnaires: printing the paper questionnaire, printing the 2 reminders to nonresponders, salary for student workers to pack all the invitations and scanning of returned paper questionnaire, postage (the invitations, prepaid envelope of returned questionnaires, and the 2 reminders to the nonresponders), 2 tablet computers (the cost of the first divided by all 800 in the Web/tablet mode, the cost of the second divided between the nonresponders in the paper, paper/Web, and Web mode), and layout of a website for

the online questionnaire. Packing of the 2 reminders and programming of the questionnaire involved no extra cost that had to be covered by the general budget at the Department of Public Health, Aarhus University. Costs in Euro (€) are stated in 2011 prices.

The Danish 2011 survey was registered at The Danish Protection Agency (Journal number: 2011-41-6230).

The Common Nordic Survey

The NordChild survey was first conducted in 1984 and then in 1996 in all the Nordic countries: Finland, Iceland, Norway, Sweden, and Denmark [16]. In each of the years the survey took place, approximately 3000 randomly selected parents of children between the ages of 2 and 17 years were invited to participate in each of the Nordic countries. In total, 46,590 children were invited during the 3 periods. Except during the Danish 2011 survey, all invited respondents were asked to fill out a paper version of the questionnaire. The questionnaire largely consisted of the same questions each time the survey was conducted. An additional 13 questions were added in the 2011 survey; thus, the questionnaires distributed in 2011 in the 5 Nordic countries consisted of 73 questions. Information on numbers of reminders, the use of an incentive, and the final response rate in each country were collected from the national contact persons.

Statistical Analysis

Descriptive analysis was performed using chi-square tests (χ^2). In the Danish 2011 survey, the responders and nonresponders were tested for each of the following descriptive characteristics, stratified by the 4 different modes: gender, age of the child, mother's and father's age, and urbanity of residence. The urbanity at municipality level was divided into 2 categories: more than and less than 100,000 citizens (the capital area of Copenhagen, Aarhus, Aalborg, Odense, Esbjerg, and Vejle in the top category). An odds ratio with 95% confidence interval (CI) estimated differences in the response rates between the 4 modes. In the analysis, the proportions of the included participants in each mode were based on initial invitation. The overall response rate for NordChild where tested in each of the years 1984, 1996, and 2011. The statistical analyses were performed in Stata version 11 (StataCorp LP, College Station, TX, USA). A *P* value <.05 was considered statistically significant.

Results

The Danish Survey

Overall, 1704 of 3148 (54.13%) respondents answered the questionnaire. The percentages of received questionnaires by days are shown in Figure 1. The highest response was seen in the Web/tablet mode in the first 20 days. The highest final response rate was obtained in the paper mode in which 443 of 789 (56.2%) responded. A similar response rate was seen in the paper/Web, Web, and Web/tablet modes: 422 of 786 (53.7%), 420 of 787 (53.4%), and 419 of 786 (53.3%) returned the questionnaire, respectively. The overall response rate before first reminder was 34.75% (1094/3148); stratified by modes, it was 35.0% (276/789) in the paper mode, 32.8% (258/786) in

the paper/Web mode, 31.6% (249/787) in the Web mode, and 39.6% (311/786) in the Web/tablet mode. For Web and Web/tablet modes that were comparable except for an incentive, the response rate was statistically significant higher in the Web/tablet mode ($P=.001$) before the first reminder. The overall response rate before second reminder was 49.78% (1567/3148); stratified by modes, it was 52.3% (413/789) in the paper mode, 47.8% (376/786) in the paper/Web mode, 48.9% (385/787) in the Web mode, and 50.0% (393/786) in the Web/tablet mode.

In connection with the first reminder, the nonresponders in the paper, paper/Web, and Web modes were also given the opportunity to win a tablet computer as a means of motivation.

The characteristics of the responders and nonresponders in the Danish 2011 survey are given in Table 1. For the age groups of the children, the highest response rates were among parents of children aged 2 to 5 years in the paper/Web (106/422, 56.7%), Web (114/420, 60.6%), and Web/tablet (100/419, 55.3%) modes. For the paper mode, the highest response rate was seen in parents of children aged 10 to 13 years (122/443, 62.6%). There were a higher number of responders from intermediate-sized urban areas outside the largest cities in Denmark.

Compared to the paper mode, the odds for response rate in the paper/Web mode decreased by 9% (OR 0.91, 95% CI 0.74-1.10)

and by 11% (OR 0.89, 95% CI 0.73-1.09) in the Web and Web/tablet mode (Table 2). Of the 422 responders in the paper/Web mode, 281 (66.6%) preferred to answer the questionnaire by paper and 141 (33.4%) preferred the Web-based version ($P<.001$). Of responders who answered the first question in the questionnaire, 413 of 419 (98.6%) responders in the paper/Web mode also answered the last question. In the paper mode, these numbers were 427 of 442 (96.7%), 387 of 416 (93.0%) in the Web/tablet mode, and 377 of 416 (90.6%) in the Web mode.

The costs per responder for distribution and collection of the paper questionnaires and the Web-based questionnaires are shown in Table 3. The cost of the paper questionnaires (€0.02) was twice that of the Web-based questionnaires (€0.55) with the postage as the single most costly item.

The Common Nordic Survey

In the common Nordic survey, the total numbers of responders declined from 10,291 of 15,339 (67.09%) in 1984 and 10,667 of 15,254 (69.93%) in 1996 to 7805 of 15,945 (48.95%) in 2011 (Table 4). For the previous years, the response rates for the NordChild 2011 were statistically significantly different across the participating countries ($P<.001$). An 8% difference was observed between the lowest rate in Sweden (1461/3197, 45.70%) and the highest in Denmark (1704/3148, 54.13%).

Figure 1. Percentage of received questionnaires by days for the paper (paper-based questionnaire only), paper/Web (paper- and/or Web-based questionnaire), Web (Web-based questionnaire only), and Web/tablet (Web-based questionnaire with tablet computer incentive within 14 days of invitation) modes in the Danish 2011 NordChild questionnaire survey.

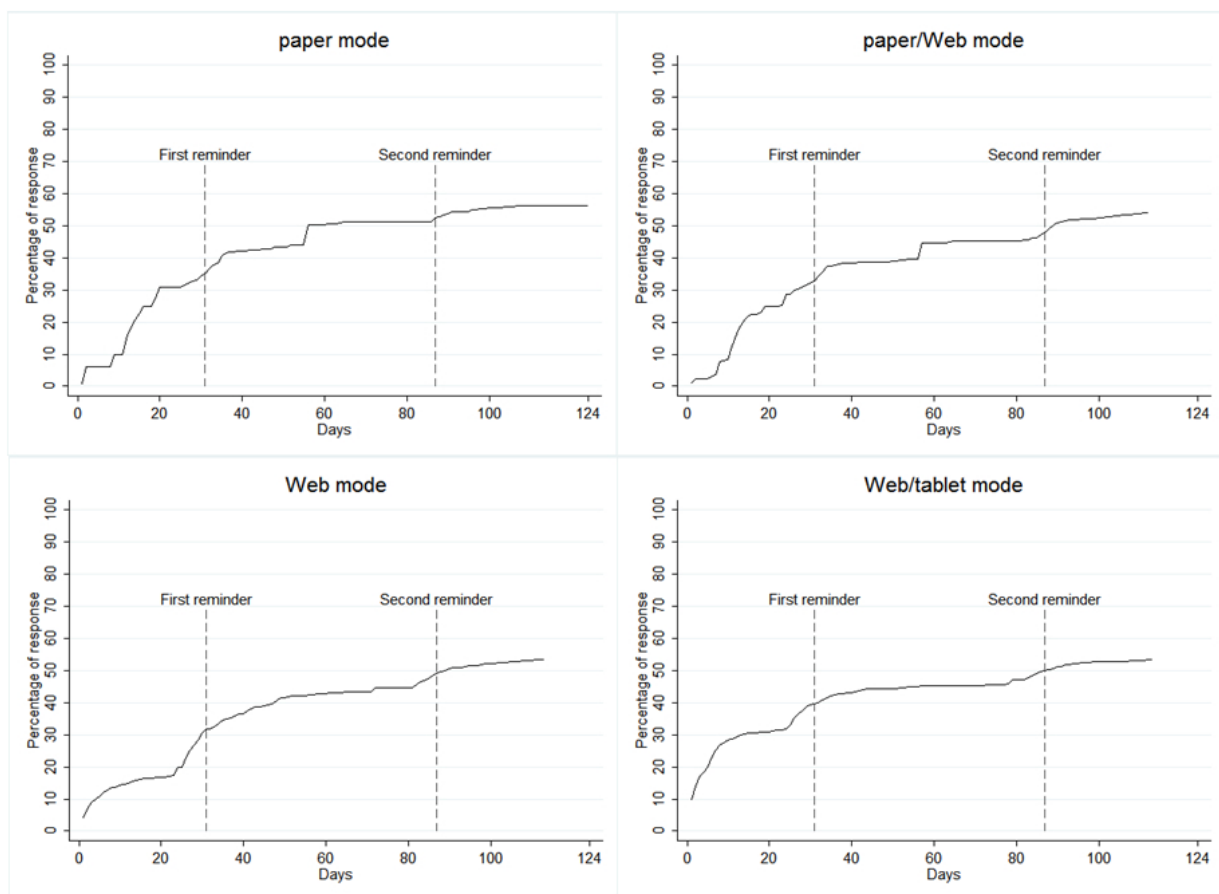


Table 1. Characteristics of the responders in the Danish part of the 2011 NordChild questionnaire survey.

Characteristics	Mode ^a											
	Paper (n=789)		Paper/Web (n=786)			Web (n=787)			Web/tablet (n=786)			
	Yes	No	<i>P</i> ^b	Yes	No	<i>P</i> ^b	Yes	No	<i>P</i> ^b	Yes	No	<i>P</i> ^b
Responding, n	443	346		422	364		420	367		419	367	
Gender, n (%)												
Male	234 (59.1)	162 (40.9)		198 (50.0)	198 (50.0)		203 (51.7)	190 (48.4)		209 (53.3)	183 (46.7)	
Female	209 (52.2)	184 (46.8)	.09	224 (57.4)	166 (42.6)	.04	217 (55.1)	117 (44.9)	.34	210 (53.3)	184 (46.7)	.99
Age of child (years), n (%)												
2-5	109 (57.1)	82 (42.9)		106 (56.7)	81 (43.3)		114 (60.6)	74 (39.4)		100 (55.3)	81 (44.7)	
6-9	111 (56.9)	84 (43.1)		111 (54.7)	92 (45.3)		108 (54.6)	90 (45.4)		103 (50.0)	103 (50.0)	
10-13	122 (62.6)	73 (37.4)		101 (50.8)	98 (49.2)		100 (50.3)	99 (49.7)		108 (53.7)	93 (46.3)	
14-17	101 (48.6)	107 (51.4)	.04	104 (52.8)	93 (47.2)	.68	98 (48.5)	104 (51.5)	.08	108 (54.6)	90 (45.4)	.73
Maternal age (years), n (%)												
<35	71 (50.3)	70 (49.7)		71 (50.7)	69 (49.3)		66 (52.8)	59 (47.2)		57 (41.9)	79 (58.1)	
35-44	265 (57.5)	196 (42.5)		237 (53.4)	207 (46.6)		246 (53.4)	215 (46.6)		248 (54.0)	211 (46.0)	
≥45	105 (58.5)	75 (41.7)	.35	112 (56.6)	86 (43.5)	.56	104 (53.1)	92 (46.9)	.99	113 (60.1)	75 (39.9)	.01
Paternal age (years), n (%)												
<35	41 (52.6)	37 (47.4)		45 (51.1)	43 (48.9)		37 (51.4)	35 (46.6)		40 (51.3)	38 (48.7)	
35-44	235 (57.3)	175 (42.7)		203 (55.5)	163 (44.5)		226 (54.1)	192 (45.9)		194 (50.1)	193 (49.9)	
≥45	160 (55.9)	126 (44.1)	.73	168 (53.5)	146 (46.5)	.73	148 (52.9)	132 (47.1)	.89	175 (59.5)	119 (40.5)	.05
Urban area, n (%)												
≤100,000	303 (58.5)	215 (41.5)		266 (53.9)	228 (46.1)		290 (55.3)	234 (44.7)		280 (54.1)	238 (45.9)	
>100,000	140 (51.7)	131 (48.3)	.07	156 (53.4)	136 (46.6)	.91	130 (49.4)	133 (50.6)	.12	139 (51.9)	129 (48.1)	.56

^aPaper: paper version of questionnaire, paper/Web: paper and/or Web-based questionnaire, Web: Web-based questionnaire, Web/tablet: Web-based questionnaire with opportunity to win a tablet computer within the first 14 days after invitation.

^bBetween responders (Yes) and nonresponders (No).

Table 2. The odds ratio of response rate by the 4 different modes in the Danish 2011 NordChild questionnaire survey.

Mode ^a	Odds ratio (95% CI)
Paper	1 (ref)
Paper/Web	0.91 (0.74, 1.10)
Web	0.89 (0.73, 1.09)
Web/tablet	0.89 (0.73, 1.09)

^aPaper: paper version of questionnaire, paper/Web: paper and/or Web-based questionnaire, Web: Web-based questionnaire, Web/tablet: Web-based questionnaire with opportunity to win a tablet computer within the first 14 days after invitation.

Table 3. Costs (2011 prices) per respondent for distribution and collection of the Danish 2011 NordChild questionnaire survey.

Cost components	Costs (€) ^a			
	Paper	Web	Paper-Web	Paper/Web
Questionnaire, print	1.39	—	—	—
Envelope				
Invitation	0.29	0.10	0.19	2.9
2× reminders	0.20	0.20	0.00	1.0
Student worker				
Packing, invitation	0.12	0.05	0.07	2.4
Scanning	0.86	—	—	-
Postage				
Invitation	2.49	0.74	1.75	3.4
Prepaid envelope	1.55	—	—	-
2× reminders	1.48	1.48	0.00	1.0
Tablet computer				
First distribution	—	0.76	—	-
Second distribution	0.40	0.31	0.09	1.3
Print of reminder				
First	0.04	0.04	0.00	1.0
Second	0.20	0.20	0.00	1.0
Layout website	—	0.67	—	-
Total per responder	9.02	4.55		
Measurement			4.47	2.0

^aPaper: paper questionnaires (paper and paper/Web), Web: Web-based questionnaires (Web and Web/tablet), paper-Web: costs of paper questionnaires minus costs of Web questionnaires, paper/Web: ratio of costs of paper questionnaires and costs of Web questionnaires, —: no cost.

Table 4. The response rate and description of data collection for the 2011 NordChild questionnaire survey.

Data collection	Participating countries					Total	P value ^a
	Denmark	Finland	Iceland	Norway	Sweden		
Responders, n (%)							
1984	2219 (73.65)	2705 (83.21)	1577 (59.51)	1856 (55.80)	1934 (62.41)	10,291 (67.09)	<.001
1996	2169 (68.64)	2384 (79.49)	2048 (68.11)	1936 (64.51)	2130 (69.00)	10,667 (69.93)	<.001
2011	1704 (54.13)	1538 (48.06)	1521 (47.53)	1581 (49.41)	1461 (45.70)	7805 (48.95)	<.001
Mode of questionnaire^b							
1984	PQ	PQ	PQ	PQ	PQ		
1996	PQ	PQ	PQ	PQ	PQ		
2011	PQ & WBQ ^c	PQ	PQ	PQ	PQ		
Number of reminders							
1984	2	2	2	2	2		
1996	2	2	2	2	2		
2011	2 ^d	1	1 ^e	2	2		
Incentive							
1984	No	No	No	No	No		
1996	No	No	No	No	No		
2011	Yes ^f	No	No	No	No		

^aP value for difference in response rates.

^bPQ: paper questionnaires, WBQ: Web-based questionnaires.

^cDistribution: 800 paper only, 800 paper or Web-based, 1600 Web-based only.

^dLog-on code to the Web-based questionnaire.

^eThank-you note to all invited participants.

^fTablet computer (2 incentives during different points of time).

Discussion

Principal Results

The main findings of this study are that 1704 of 3148 (54.13%) invited respondents answered the Danish questionnaire. In the 4 modes, the response rate was slightly higher, but not statistically significant, in the paper (443/789, 56.2%) mode than in the paper/Web (422/786, 53.7%), Web (420/787, 53.3%), and the Web/tablet (419/786, 53.3%) modes. The failure of a nonmonetary incentive to affect response rates in all 4 modes investigated in the Danish part of the study may partly be because of a lack of a comparable reference group. The Web-based questionnaire mode carried fewer costs than the other modes. Overall, 7805 of 15,945 (48.95%) participated in the NordChild 2011 with statistically significant differences in response rates between the countries ($P < .001$).

Comparison With Previous Studies

The slightly higher final response rate in the paper mode than in the other modes confirms findings reported in some other studies [5,7], but not in all [8]. The latter study reported a higher response rate in the Web-based questionnaire than in the paper-based version, but its study population was younger (mean age 30 years) than the population in our study in which the mean

parental age was approximately 40 years. In the studies which reported the highest response rates for the paper version, the study populations' age range was either equivalent to the age span of our population (30-60 years) [7] or it was slightly higher (mean age 55 years) [5].

The study with the younger study population compared a Web-based version with a mixed-mode questionnaire [8]. It showed that among participants who had a choice between a paper-based and a Web-based questionnaire, most preferred to answer the questionnaire using the paper version (83%), whereas only 17% answered the questionnaire online [8]. This inclination was also seen in the present study in which two-thirds of the respondents in the paper/Web mode returned a paper version of the questionnaire.

Questionnaire participants' propensity to respond is likely shaped by the relevance they ascribe to the questionnaire. Hence, the participants in previous studies [5,7,8] were all current or previous victims of disease or a benign abnormality and their high response rates (60%-89%) may testify to the relevance of the questionnaires to their current or previous health. It is possible that a survey on health and welfare among children and young people is of little concern to some parents; therefore, we may have overvalued it as a clear topic and motivation to participate in our study.

Assessing the effect of the chance that a single responder would receive a large nonmonetary incentive was only possible for the 2 comparable modes, Web and Web/tablet, in the period from the initial invitation to the first reminder. Therefore, the present study is only contributing with information to the lacking knowledge about the effect on the response rates of offering a large, nonmonetary incentive in a Web-based questionnaire survey in the first month, conducted among a population familiar with and enjoying largely unrestrained access to the Internet. The increase in received questionnaires between the first and the second reminders for the paper and Web mode of 17.3% and 15.0% for the paper/Web mode partly may be ascribed to the incentive and partly because of the effect of being reminded about the questionnaire. An alternative to the incentive could have been to offer to communicate the study results to the participants, but previous studies have shown that this does not increase the response rates [9,21].

We found the costs for the paper questionnaire to be double the costs of the Web-based questionnaire, which is in-line with other studies [5,6]. The lower costs and the advantages of the Web logistics suggest that the Web-based questionnaire may be an alternative to paper questionnaires. The slightly lower response rate means that the costs per respondent may increase [1], which implies that the difference in costs between the paper and Web-based questionnaire may have been even larger if we had obtained a higher participation rate.

Strengths and Limitations

A strength of the present study is that the invited study population is clearly representative of the whole population. Furthermore, we have access to registers and, thereby, the possibility to make follow-up studies of the children's health and welfare by use of the unique 10-digit personal identification number. Our study also has certain limitations. The Danish data are generalized to the Nordic countries and to other countries with similar access to the Internet, but the Nordic data may not be generalized to countries with unequal access to the Internet because of different Internet patterns of behavior. Furthermore, our data may not be extrapolated to other age groups because of possible changes in Internet patterns of behavior.

Our anticipation was not met that the Web-based questionnaires would feature the highest final response rate and that the final response rate would be highest in the Web/tablet mode, although the highest response rate before the first reminder was seen in the Web/tablet mode. This may be because of lack of a comparable reference group because all participants approached for the paper, paper/Web, Web, and Web/tablet modes were offered a chance to win a tablet computer during different points of time. It may also be because of the length of the questionnaire with 73 questions. A questionnaire of this length requires that a relatively large amount of time is spent answering the questionnaire. The length of the questionnaire was due primarily to addition of the Strength and Difficulties Questionnaire [22] to the NordChild 2011 survey, which had a considerable number of subquestions. The paper version was the slightly preferred questionnaire mode based on the difference in response rate and that most respondents in the paper/Web mode returned the paper version of the questionnaires. This preference for the paper

version may be rooted in tradition and in convenience answering a paper questionnaire in several rounds. The 12-character log-in code to the Web-based questionnaire may have hindered the process of completing it in several rounds. However, there may also be a limit to how long a Web-based questionnaire can be. Yet, a certain number of questions must be asked in health questionnaire surveys to allow due analysis of the issues explored. Decreasing the number of questions in the present study would preclude comparison with the previous studies in 1984 and 1996, which would rob the NordChild survey of its potential to make comparisons over time.

We first planned to offer only 1 tablet computer to those who were invited in the Web/tablet mode. However, an overall response rate of 34.8% 1 month after the initial invitation was unacceptably low. Therefore, we offered another tablet computer to motivate the nonresponders in the paper, paper/Web, and Web mode with the mailed first reminder. The effect of this initiative may be blurred by allowing the nonresponders in the paper mode to answer the questionnaire online as well. The effect of offering an incentive should also be appraised in light of the nature of that incentive. Winning a tablet computer might simply not have been sufficiently attractive to make potential respondents engage in answering a comprehensive questionnaire because Denmark saw a significant rise in household possession of tablet or minicomputers from 9% in 2011 to 19% in 2012 (ie, during the study period) [23].

From a socioeconomic perspective, it would have been interesting to know which responders sent in their questionnaires right away, and if there were social differences in the need for motivation rooted in the hope of winning a tablet computer. Although we registered the date the questionnaires were returned, we collected no information about what had motivated the responders to answer the questionnaire.

The Total Response Rate in the NordChild survey

The highest response rate in the NordChild 2011 survey was obtained in the Danish part. For the NordChild surveys in 1984 and 1996, the highest response rate was observed in Finland. The responders and nonresponders may differ between the Nordic countries, but response rates varied less between the Nordic countries in the NordChild 2011 survey than in the surveys in 1984 and 1996. In general, the data collection for the NordChild 2011 survey was less consistent than the data collection in 1984 and 1996. Thus, in Denmark, an online questionnaire was introduced together with an incentive, whereas 1 reminder was mailed in Iceland and Finland compared with 2 reminders in Sweden, Norway, and Denmark. Furthermore, a thank-you note was sent to all the invited respondents in Iceland, including 1 sentence as a reminder if the sampled person had not yet responded. The proportion of people who had a priori declined research participation in the form of postal surveys also varied between the Nordic countries. In Denmark, this figure exceeded 10%, whereas it was approximately 4% in Finland. In other Nordic countries, the survey was also sent to those who had banned the commercial use of their addresses. In other words, it remains unknown how much these differences biased the results.

Despite statistically significant differences in response rates, the different starting points of distribution, and the prospect of winning a tablet computer in the Danish part of the survey, the clinical response rate for the NordChild 2011 study was almost the same across the Nordic countries. The participants in the other 4 Nordic countries were unaware of the incentive in the Danish part of the NordChild 2011 survey which entails nondifferential selection bias. We believe that the low response rate observed in all the Nordic countries is more likely because of a general decrease in volunteerism, higher demands for participation, and oversurveying [1].

Perspectives

The declining response rates in population-based surveys in general are a challenge to epidemiology. First, the overall response rate of 48.95% in the NordChild 2011 study makes it difficult to compare the results of the survey with the results of previous NordChild surveys in which the response rates were considerably higher. Second, the low participation rate could raise the question whether the responders are representative of the general population. Analyses linking data with administrative registers are needed to determine to which extent randomly selected responders represent the general population. Furthermore, information about what motivates parents to participate in a scientific study about their child's health also needs to be elucidated to determine to which extent incentives, such as the results of the survey, should be offered to raise response levels to desired levels.

The present study suggests that Web-based questionnaires and the chance of winning a tablet computer are unlikely to solve the problem of low respondent attendance; therefore, other

solutions must be considered. Like many other developed countries, the Nordic countries regularly conduct routine child health examinations from gestation to the end of school age that, together with the school setting itself, provide a platform for gathering valid and high-coverage survey information. By using a clinical meeting with the child and the mother or father (when the child is younger) could be a promising way to get more complete epidemiological data at the population level. In Denmark, attendance to these routine examinations is in the range 80% to 95% and the development of common Web-based tools with an interactive clinical component, such as Schoolhealth.eu, may be a means for monitoring the development of child health [24]. Future Web-based surveys may also benefit from the use of smartphones, which every second Dane is in possession of now [25].

Conclusions

Web-based questionnaires could replace traditional paper questionnaires based on comparable response rates and lower costs. The increasing effect on the response rate on participants replying for a nonmonetary incentive could only be estimated within the 2 Web-based questionnaire modes before the first reminder. The difference in costs between the paper and Web-based questionnaires will favor the Web-based mode provided high response rates can be obtained. Web-based questionnaires provide an alternative to the traditional paper version; however, we need to consider alternative platforms to reach higher participation rates in such population surveys. Such alternatives should reflect the development of electronic devices and the ways in which the population primarily accesses the Internet.

Acknowledgments

This study was funded by Trygfonden grant 7-11-00-1396.

We thank Ynna Margot Nielsen, Section of General Practice, Department of Public Health at Aarhus University, who helped with layout of the paper version of the Danish questionnaire. We also thank Søren Beck Morsing and Marianne Pedersen, Data Management, Department of Public Health at Aarhus University, for their effort with the layout of the Web-based questionnaire and data cleaning.

We also thank Jenný Ingudóttir, Jon Oskar Gudlaugsson, and Geir Gunnlaugsson at the Directorate of Health in Reykjavik, Iceland, and Heljä-Marja Surcel, THL National Institute for Health and Welfare in Oulu, Finland, as project manager for the 2011 data collection in Finland.

Conflicts of Interest

The PhD project for Lena Hohwü is fully funded with a grant from Aarhus University, Denmark.

References

1. 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/)]
2. Morton LM, Cahill J, Hartge P. Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol* 2006 Feb 1;163(3):197-203 [FREE Full text] [doi: [10.1093/aje/kwj036](https://doi.org/10.1093/aje/kwj036)] [Medline: [16339049](https://pubmed.ncbi.nlm.nih.gov/16339049/)]
3. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology* 2006 Jul;17(4):413-418. [doi: [10.1097/01.ede.0000220549.14177.60](https://doi.org/10.1097/01.ede.0000220549.14177.60)] [Medline: [16755269](https://pubmed.ncbi.nlm.nih.gov/16755269/)]
4. Griffin JM, Simon AB, Hulbert E, Stevenson J, Grill JP, Noorbaloochi S, et al. A comparison of small monetary incentives to convert survey non-respondents: a randomized control trial. *BMC Med Res Methodol* 2011;11:81 [FREE Full text] [doi: [10.1186/1471-2288-11-81](https://doi.org/10.1186/1471-2288-11-81)] [Medline: [21615955](https://pubmed.ncbi.nlm.nih.gov/21615955/)]

5. Zuidgeest M, Hendriks M, Koopman L, Spreeuwenberg P, Rademakers J. A comparison of a postal survey and mixed-mode survey using a questionnaire on patients' experiences with breast care. *J Med Internet Res* 2011;13(3):e68 [FREE Full text] [doi: [10.2196/jmir.1241](https://doi.org/10.2196/jmir.1241)] [Medline: [21946048](https://pubmed.ncbi.nlm.nih.gov/21946048/)]
6. Scott A, Jeon SH, Joyce CM, Humphreys JS, Kalb G, Witt J, et al. A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors. *BMC Med Res Methodol* 2011;11:126 [FREE Full text] [doi: [10.1186/1471-2288-11-126](https://doi.org/10.1186/1471-2288-11-126)] [Medline: [21888678](https://pubmed.ncbi.nlm.nih.gov/21888678/)]
7. 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/)]
8. van den Berg MH, Overbeek A, van der Pal HJ, Versluys AB, Bresters D, van Leeuwen FE, et al. Using web-based and paper-based questionnaires for collecting data on fertility issues among female childhood cancer survivors: differences in response characteristics. *J Med Internet Res* 2011;13(3):e76 [FREE Full text] [doi: [10.2196/jmir.1707](https://doi.org/10.2196/jmir.1707)] [Medline: [21955527](https://pubmed.ncbi.nlm.nih.gov/21955527/)]
9. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009(3):MR000008. [doi: [10.1002/14651858.MR000008.pub4](https://doi.org/10.1002/14651858.MR000008.pub4)] [Medline: [19588449](https://pubmed.ncbi.nlm.nih.gov/19588449/)]
10. Wilson PM, Petticrew M, Calnan M, Nazareth I. Effects of a financial incentive on health researchers' response to an online survey: a randomized controlled trial. *J Med Internet Res* 2010;12(2):e13 [FREE Full text] [doi: [10.2196/jmir.1251](https://doi.org/10.2196/jmir.1251)] [Medline: [20457556](https://pubmed.ncbi.nlm.nih.gov/20457556/)]
11. Gates S, Williams MA, Withers E, Williamson E, Mt-Isa S, Lamb SE. Does a monetary incentive improve the response to a postal questionnaire in a randomised controlled trial? The MINT incentive study. *Trials* 2009;10:44 [FREE Full text] [doi: [10.1186/1745-6215-10-44](https://doi.org/10.1186/1745-6215-10-44)] [Medline: [19545427](https://pubmed.ncbi.nlm.nih.gov/19545427/)]
12. Ulrich CM, Danis M, Koziol D, Garrett-Mayer E, Hubbard R, Grady C. Does it pay to pay? A randomized trial of prepaid financial incentives and lottery incentives in surveys of nonphysician healthcare professionals. *Nurs Res* 2005 Jun;54(3):178-183. [Medline: [15897793](https://pubmed.ncbi.nlm.nih.gov/15897793/)]
13. Wyrick DL, Bond L. Reducing sensitive survey response bias in research on adolescents: a comparison of web-based and paper-and-pencil administration. *Am J Health Promot* 2011 Jun;25(5):349-352. [doi: [10.4278/ajhp.080611-ARB-90](https://doi.org/10.4278/ajhp.080611-ARB-90)] [Medline: [21534838](https://pubmed.ncbi.nlm.nih.gov/21534838/)]
14. Olsen F, Abelsen B, Olsen JA. Improving response rate and quality of survey data with a scratch lottery ticket incentive. *BMC Med Res Methodol* 2012;12(1):52 [FREE Full text] [doi: [10.1186/1471-2288-12-52](https://doi.org/10.1186/1471-2288-12-52)] [Medline: [22515335](https://pubmed.ncbi.nlm.nih.gov/22515335/)]
15. Statistics Denmark. The use of ICT in the population 2012 [IT-anvendelse i befolkningen 2012]. Copenhagen, Denmark: Statistics Denmark; 2012. URL: <http://www.dst.dk/pukora/epub/upload/17443/itanv.pdf> [accessed 2013-01-02] [WebCite Cache ID 6DNDz7NM0]
16. Pedersen CR, Holstein BE, Köhler L. Parents' labour market participation as predictor of children's well-being: changes from 1984 to 1996 in the Nordic countries. *Eur J Public Health* 2005 Aug;15(4):431-436 [FREE Full text] [doi: [10.1093/eurpub/cki012](https://doi.org/10.1093/eurpub/cki012)] [Medline: [16155067](https://pubmed.ncbi.nlm.nih.gov/16155067/)]
17. Pedersen CB. The Danish Civil Registration System. *Scand J Public Health* 2011 Jul;39(7 Suppl):22-25. [doi: [10.1177/1403494810387965](https://doi.org/10.1177/1403494810387965)] [Medline: [21775345](https://pubmed.ncbi.nlm.nih.gov/21775345/)]
18. Statistics Denmark. Memorandum of individual unregistration to statistical and scientific studies 2008 [Notat om Forskerbeskyttelse i CPR 2008]. Copenhagen, Denmark: Statistics Denmark; 2008. URL: http://www.dst.dk/~media/Sc4/Statistik/dokumentation/metode/Forsker2008/notat_om_forskerbeskyttelse_2008-pdf.pdf [accessed 2013-01-16] [WebCite Cache ID 6DiWrXpN6]
19. Statistics Denmark. Prohibiting delivery of individual unregistration to statistical and scientific studies 2009 [Forskerbeskyttelse 2009]. Copenhagen, Denmark: Statistics Denmark; 2009. URL: <http://www.dst.dk/da/TilSalg/interview/Kunde/Forskerbeskyttelse.aspx> [accessed 2013-01-16] [WebCite Cache ID 6DiXQ1z7I]
20. SurveyXact. Knowledge is the avenue to results - SurveyXact is Your direct avenue to knowledge [Viden er vejen til resultater - SurveyXact er din direkte vej til viden] URL: <http://www.surveyxact.dk/loesninger> [accessed 2013-04-30] [WebCite Cache ID 6GGNwU8c8]
21. Cockayne S, Torgerson DJ. A randomised controlled trial to assess the effectiveness of offering study results as an incentive to increase response rates to postal questionnaires [ISRCTN26118436]. *BMC Med Res Methodol* 2005;5:34 [FREE Full text] [doi: [10.1186/1471-2288-5-34](https://doi.org/10.1186/1471-2288-5-34)] [Medline: [16250910](https://pubmed.ncbi.nlm.nih.gov/16250910/)]
22. Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiatry* 1999 Jul;40(5):791-799. [Medline: [10433412](https://pubmed.ncbi.nlm.nih.gov/10433412/)]
23. Statistics Denmark. The families possession of home appliances by type of consumption [Familiernes besiddelse af elektronik i hjemmet efter forbrugsart]. Copenhagen, Denmark: Statistics Denmark URL: <http://www.dst.dk/da/search.aspx?q=pc+tablet> [accessed 2013-01-16] [WebCite Cache ID 6DiXkWWkp]
24. Skolesundhed.dk. An interactive tool developed in Denmark for children's health examinations during school age URL: <http://schoolhealth.eu/> [accessed 2013-01-16] [WebCite Cache ID 6DiYPicVO]

25. Statistics Denmark. Every second Dane takes access to the Internet using a mobile phone [Hver anden dansker går på internettet med mobilen]. Copenhagen, Denmark: Statistics Denmark; 2012 Jul 20. URL: <http://www.dst.dk/pukora/epub/Nyt/2012/NR376.pdf> [accessed 2013-01-16] [WebCite Cache ID 6DiYklCRy]

Abbreviations

NordChild: “Health and welfare among children and young people in the Nordic countries” questionnaire survey

PQ: all paper versions of the questionnaire (paper and paper/Web)

paper: paper questionnaire only

paper/Web: paper questionnaire as well as log-in code to the Web-based questionnaire

WBQ: all Web-based questionnaires (Web and Web/tablet)

Web: Web-based questionnaire only

Web/tablet: Web-based questionnaire plus an incentive consisting of a chance to win a tablet computer

Edited by G Eysenbach; submitted 12.03.13; peer-reviewed by M van Gelder, P Edwards, S Lim Choi Keung, X Zhang; comments to author 24.04.13; revised version received 24.05.13; accepted 25.06.13; published 26.08.13.

Please cite as:

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

URL: <http://www.jmir.org/2013/8/e173/>

doi: [10.2196/jmir.2595](https://doi.org/10.2196/jmir.2595)

PMID: [23978658](https://pubmed.ncbi.nlm.nih.gov/23978658/)

©Lena Hohwü, Heidi Lyshol, Mika Gissler, Stefan Hrafn Jonsson, Max Petzold, Carsten Obel. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.08.2013. 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

Using the Internet to Help With Diet, Weight, and Physical Activity: Results From the Health Information National Trends Survey (HINTS)

Scout N McCully¹, BA; Brian P Don¹, MA; John A Updegraff¹, PhD

Department of Psychology, Kent State University, Kent, OH, United States

Corresponding Author:

John A Updegraff, PhD
Department of Psychology
Kent State University
321 Kent Hall
Kent, OH, 44242-0001
United States
Phone: 1 330 672 4731
Fax: 1 330 672 3786
Email: jupdegr1@kent.edu

Abstract

Background: The Internet offers a viable platform for cost-effective and wide-reaching health interventions. However, little is known about use of the Internet to help with diet, weight, and physical activity (DWPA) using a nationally representative sample from the United States.

Objective: To (1) assess the demographic characteristics of people who use the Internet to help with DWPA, (2) assess whether usage trends changed over time, and (3) investigate the associations between using the Internet for DWPA and health behaviors.

Methods: Data on Internet users from the 2007 and 2011 iterations of the Health Information National Trends Survey (HINTS), N=4827 were analyzed using multiple logistic regression to determine the demographic correlates of using the Internet for help with DWPA. Multiple linear regression was used to test the associations between Internet use for DWPA and three health behaviors: fruit intake, vegetable intake, and physical activity.

Results: A larger percentage of Internet users used the Internet for DWPA in 2011 (42.83%) than in 2007 (40.43%). In general, Internet users who were younger (OR 0.98, $P<.001$), more educated (OR 1.40, $P<.001$), married (OR 1.06, $P=.03$), of a minority race (non-Hispanic blacks: OR 1.14, $P=.02$; Hispanics: OR 1.42, $P=.01$), and who had a higher Body Mass Index (BMI) (OR 1.04, $P<.001$) were more likely to use the Internet for DWPA. Across survey years, gender was not associated with using the Internet for DWPA (OR 1.03, $P=.12$), but there was a significant interaction between survey year and gender (OR 1.95, $P=.002$); in 2007, men were more likely to use the Internet for DWPA, but women were more likely to do so in 2011. Using the Internet for DWPA was associated with more vegetable intake ($B=.22$, $P=.002$), more fruit intake ($B=.19$, $P=.001$), and more moderate exercise ($B=.25$, $P=.001$), although the strength of the associations between using the Internet for DWPA and fruit intake and exercise was weaker in 2011 than in 2007.

Conclusions: Contrary to prior research, our population-level study did not show a pronounced gender difference in the use of the Internet for DWPA. Our results support the increasing viability of the Internet as a platform for behavior change intervention, as a growing percentage of Internet users are turning to the Internet for help with DWPA. Additionally, using the Internet for DWPA is associated with better DWPA-related health behaviors.

(*J Med Internet Res* 2013;15(8):e148) doi:[10.2196/jmir.2612](https://doi.org/10.2196/jmir.2612)

KEYWORDS

physical activity; weight loss; dietary habits; Web-based; Internet; utilization

Introduction

Poor nutrition and physical inactivity have significant and negative implications for individuals and society at large. On the individual level, poor nutrition and physical inactivity are risk factors for obesity, type 2 diabetes, and heart disease [1-4]. On the societal level, unhealthy lifestyles create a heavy economic burden through largely preventable diseases [5,6]. Thus, focus has turned to prevention, such as modification of behavioral risk factors to reduce incidence of disease. Many evidence-based clinical interventions have been developed to help people manage their diet, weight, and physical activity (DWPA), and a plethora of open-access and/or commercial DWPA programs are available via the Internet [7-9]. Internet-delivered programs are unique in their ability to cost-effectively reach large numbers of users across geographically dispersed areas, provide anonymity for users who wish it, and reduce time and travel demands that in-person programs necessitate [10,11]. Despite the burgeoning availability of DWPA programs, few studies have examined their use at a population level.

The demographic trends in DWPA use provide essential information to those seeking to develop, implement, and evaluate Internet-based DWPA programs. Demographic information provides a basis for tailored and targeted programs, which produce better health outcomes than nontailored or nontargeted programs (see [12] for review). Yet little is currently known about who uses Internet-based DWPA programs because extant usership statistics are based on self-selected samples from randomized clinical trials (RCTs) and certain commercially offered or open-access programs [13-16]. Existing studies suggest that those who elect to participate in Internet-based RCTs for diet and weight loss are primarily female, middle aged, and college educated [13-15]. Specific commercially offered or open-access programs for DWPA indicate a similar pattern; for example, a study of the commercial Web-based weight loss program, *The Biggest Loser Club, Australia*, revealed that 86% of the nearly 10,000 enrollees were female, the majority was of moderate to high socioeconomic status, and the mean age was 35.7 years [16]. Yet, because these statistics describe samples of individuals openly seeking enrollment in an official trial or open-access weight loss program, they may not reflect the average demographics of ad libitum usership of the Internet for DWPA, as only a minority of people who visit a website with an RCT enrollment opportunity elect to enroll [15]. Thus, examining self-selected enrollers in clinical trials or users of a select few open-access and commercial programs may not provide complete demographic data of national usage. Therefore, the first aim of the current study is to document the demographic profile of those who use the Internet to help with DWPA using a nationally representative sample.

Regardless of who uses Internet-based DWPA programs, the viability of the Internet as a platform for wide-reaching health behavior interventions is dependent on these programs actually reaching a large population of users. However, because no large nationally representative studies on trends in Internet use for DWPA have been conducted, it is not known whether Internet use for DWPA is increasing or decreasing. Therefore, the second

aim of our study is to investigate temporal trends in usership—namely, whether a higher percentage of Internet users used the Internet for DWPA in 2011 than in 2007, and whether there are important demographic differences in usership over time. Research by the National Telecommunications and Information Administration indicates that overall use and access to the Internet is increasing; that is, more people had Internet in their homes in 2010 compared to 2007 and more people had broadband, indicating faster connection speeds [17]. Additionally, Pew Research Center polls indicate that 81% of adults used the Internet in 2012 compared to 71% in 2007 [18]. With a broader base of Internet users in general, examining whether any demographic shifts in users of the Internet for DWPA occurred between 2007 and 2011 is important, as this information may inform prospective tailoring or targeting of future DWPA programs. Thus, we predict that a higher percentage of Internet users would be using the Internet for DWPA in 2011 than in 2007. We also examine any changes over time in the demographics of people who use the Internet for DWPA.

A third aim of our study is to examine how use of the Internet for DWPA relates to adherence to DWPA-related behaviors. To date, no studies have assessed the relationship between ad libitum use of the Internet for DWPA and health behaviors in a nationally representative study. Because the viability of Internet-based programs depends on their actual association with health behaviors, we aim to examine whether use of Internet-based DWPA programs is associated with the key health behaviors of vegetable intake, fruit intake, and exercise. To investigate these relationships, we examined data from the 2007 and 2011 iterations of the US National Cancer Institute's Health Information National Trends Survey (HINTS).

Methods

Data Source

This study used data from two iterations of HINTS, administered in 2007 and 2011. HINTS is a national probability survey of US adults that assesses usage and trends in health information access and understanding. By oversampling high minority areas, HINTS provides greater precision of estimates for minority subpopulations. HINTS has been administered iteratively, and publicly accessible datasets from 2003, 2005, 2007, and 2011 are available at the HINTS website [19], along with general methodological information about the HINTS survey. The 2007 and 2011 iterations both included an item assessing Internet use for DWPA; no prior iterations included this item. In 2007, two methods were used for data collection: a random digit dial telephone survey and a paper and pencil survey. The mailed survey, but not the random digit dial survey, was used because it assessed fruit and vegetable intake on the same scale (cups per day) as the 2011 survey. The 2007 household response rate for mailed surveys was 40.0%, and the 2011 household response rates were 37.9% for the next-birthday selection method (only the adult whose next birthday is soonest completes the survey) and 35.3% for the all adult selection method (all adults in household complete the survey). More information can be found in the respective cycles' methodology reports [20,21].

Participants

A total of 3582 individuals completed the mailed survey in the 2007 HINTS, and 3959 individuals completed the 2011 survey. Of all possible participants, those who had missing data on using the Internet for DWPA (32.8%) or key demographic variables (2.5%) were excluded, resulting in a sample size of 4827 in demographic analyses. Those who had data for using the Internet for DWPA but not for health behavior variables were excluded on an analysis-by-analysis basis, with the greatest missing data for physical activity (10.5%), and negligible missingness for vegetable intake (5 participants), and fruit intake (2 participants).

Measures

Demographics

We used participants' self-reports of age, sex, level of education, and height and weight. We converted height and weight to body mass index (BMI) scores in which higher scores indicate a generally less healthy weight-to-height ratio.

Internet Use for DWPA

Participants answered one question, "In the last 12 months, have you used the Internet to: Use a website to help you with your diet, weight, or physical activity?" with a yes/no response.

Fruit and Vegetable Intake

Fruit and vegetable intake were each assessed with the question, "About how many cups of fruit [vegetables] (including 100% pure fruit [vegetable] juice) do you eat or drink each day?" Examples of 1 cup of fruits and vegetables were provided, such as "1 small apple" or "3 broccoli spears." Participants had 7 response options from "none" to "4 cups or more" such that higher scores represent greater intake. Fruit and vegetable intake scores were analyzed separately.

Physical Activity

Physical activity was assessed with the question, "In a typical week, how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace?" The 8 response options ranged from "none" to "7 days per week."

Statistical Analyses

Combining the Datasets

To analyze differences over time, the 2007 and 2011 datasets were combined using methods employed in prior HINTS analyses [22]. We modified the procedure for the 2007 and 2011 datasets, and for 2007, we used the mail-only final and replicate

weights to accurately weight the data based on our exclusive use of mail surveys.

Analytic Procedure

Multiple logistic regression was used to determine the demographic correlates of usage of the Internet for DWPA and to assess whether changes occurred in these demographic associations over time. Interaction terms between demographic variables and survey year were included in the model to identify significant changes in demographic makeup of users over time. Multiple linear regression, controlling for demographic variables, was used to analyze the strength of the relationship between use of the Internet for DWPA and the health behaviors of fruit and vegetable intake and physical activity. Interaction terms between Internet use for DWPA and survey year were added to the regressions to assess changes across time. Goodness of fit for all logistic models was assessed with the Hosmer-Lemeshow (H) test statistic and Tjur's [23] coefficient of discrimination (D). A nonsignificant H statistic indicates good fit, or minimal deviation between observed and predicted values. The D ranges from 0 (no discriminatory power) to 1 (perfect discriminatory power) and can be interpreted roughly as the percent shift in predicted versus observed probabilities compared to a null model. All statistical analyses were conducted using Stata 12, and a cutoff of $P < .05$ was used to determine statistical significance for all analyses.

Results

Demographic Predictors of Using the Internet for DWPA

Age, sex, level of education, BMI, race/ethnicity, and marital status were examined as demographic predictors of using the Internet for DWPA. Each increase in level of education (OR 1.40, 95% CI 1.38-1.42, $P < .001$) and in BMI (OR 1.04, 95% CI 1.03-1.04, $P < .001$) was associated with a significantly greater likelihood of using the Internet for DWPA, but each additional year of age was associated with significantly lower likelihood of using the Internet for DWPA (OR 0.98, 95% CI 0.97-0.98, $P < .001$). Married individuals were more likely than unmarried individuals to have used the Internet for DWPA (OR 1.06, 95% CI 1.02-1.11, $P = .03$). Both non-Hispanic blacks (OR 1.14, 95% CI 1.06-1.23, $P = .02$) and Hispanics (OR 1.42, 95% CI 1.20-1.68, $P = .01$) were more likely than non-Hispanic whites to have used the Internet for DWPA. Gender was not related to using the Internet for DWPA (OR 1.03, 95% CI 0.98-1.09, $P = .12$). See Tables 1 and 2 for unweighted group sizes, population-weighted percentages, and means and standard deviations. The logistic model provided adequate fit, $H = 3.87$, $P = .18$, $D = .065$.

Table 1. Demographic correlates of using the Internet for help with DWPA in 2007 and 2011.

	Did not use Internet for DWPA				Used Internet for DWPA			
	2007		2011		2007		2011	
	n	%	n	%	n	%	n	%
Total sample	1462	59.6	1457	57.2	961	40.4	947	42.8
Gender								
Male	878	29.9	630	29.2	640	22.7	309	19.0
Female	584	29.3	827	28.0	321	17.8	638	23.9
Race/Ethnicity								
Non-Hispanic white	1108	47.8	1006	45.6	674	30.3	622	31.2
Non-Hispanic black	137	5.2	193	6.7	113	4.7	124	4.7
Hispanic	114	6.8	113	5.6	90	5.1	100	6.3
Education								
Less than High school	60	4.6	58	3.9	22	1.4	28	2.6
High school graduate	292	14.1	235	11.5	105	5.9	93	5.7
Some college	509	24.1	491	21.0	334	17.8	282	14.0
Bachelor's	359	10.1	378	11.7	303	9.9	346	14.0
Post bacc.	242	6.8	295	9.2	197	5.4	198	7.5
Marital status^a								
Married	848	35.1	794	33.9	564	24.8	495	22.5
Single	511	24.2	518	23.9	313	16.0	351	19.7

^aMarital status was collapsed for simplicity into two categories: married (married or living as married) and single (single, divorced, separated, widowed).

Table 2. Demographic and health behavior correlates of using the Internet for help with DWPA in 2007 and 2011.

	Did not use Internet for DWPA				Used Internet for DWPA			
	2007		2011		2007		2011	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	50.30	16.12	52.74	15.32	44.64	13.44	45.72	13.99
BMI	27.32	6.27	27.40	6.28	28.14	6.60	28.51	6.74
Health behaviors^a								
Fruit intake	2.42	1.38	2.38	1.36	2.72	1.36	2.54	1.36
Vegetable intake	2.77	1.31	2.67	1.36	3.03	1.33	2.88	1.32
Moderate exercise	3.35	1.84	2.68	2.21	3.57	1.75	2.81	2.06

^aFruit (N=4816) and vegetable (N=4813) intake reflect cups per day; moderate exercise (N=4062) reflects days per week.

Changes in Use of the Internet for DWPA

To determine whether more people used the Internet for DWPA in 2011 than in 2007, we conducted a multiple logistic regression including all of the demographic variables (age, gender, BMI, education, race, and marital status) and survey year. There was a trend that Internet users in 2011 were more likely to have used the Internet for DWPA than were Internet users in 2007 (OR 1.05, 95% CI 0.99-1.12, $P=.07$). The logistic model provided adequate fit, $H=.46$, $P=.57$, $D=.065$.

To assess whether the makeup of users had changed between 2007 and 2011, we ran separate multiple logistic regressions in

which the demographic variables, survey year, and an interaction term between the survey year and the demographic variable of interest were entered. There was no change across years in BMI (OR 1.00, 95% CI 0.99-1.01, $P=.98$). However, significant changes between 2007 and 2011 were found for gender (OR 1.95, 95% CI 1.73-2.19, $P=.002$), age (OR 0.99, 95% CI 0.98-0.99, $P=.003$), education (OR 0.91, 95% CI 0.90-.93, $P=.001$), and marital status (OR 0.76, 95% CI 0.70-0.84, $P=.007$). In other words, users of the Internet for DWPA were younger, less educated, and more likely to be female and single in 2011 than in 2007. Additionally, a lower proportion of non-Hispanic blacks (OR 0.74, 95% CI 0.66-0.83, $P=.008$) and a higher proportion of Hispanics (OR 1.38, 95% CI 1.06-1.81,

$P=.04$) used the Internet for DWPA in 2011 than in 2007, compared to non-Hispanic whites. All models fit the data adequately ($P>.17$), and addition of these interactions improved model fit ($D=.074$).

Health Behaviors Associated With Using the Internet for DWPA

In general, people who used the Internet for DWPA in 2011 reported 2.54 cups of fruit intake per day, 2.88 cups of vegetable intake per day, and 2.81 days of exercise per week (see [Table 2](#)). Because we analyzed users of the Internet for DWPA only among those without missing data on the key variable, we descriptively checked health behavior means among those who did not use the Internet at all. In general, people who did not use the Internet at all reported 2.27 cups of fruit intake per day, 2.57 cups of vegetable intake per day, and 2.18 days of exercise per week. The lower levels in health behaviors among non-Internet users suggest that our analysis of Internet users was a more stringent test of the relationship between using the Internet for DWPA and health behavior.

We tested the associations between health behaviors and use of the Internet for DWPA using 3 multiple linear regressions, each controlling for all demographic variables (BMI, age, gender, education, race, and marital status) as well as survey year. People who used the Internet for DWPA reported more vegetable intake ($\beta=.08$, $B=.22$, 95% CI 0.18-0.27, $P=.002$), more fruit intake ($\beta=.07$, $B=.19$, 95% CI 0.17-0.21, $P=.001$), and more moderate exercise ($\beta=.06$, $B=.25$, 95% CI 0.22-0.29, $P=.001$) than those who did not use the Internet for DWPA.

Post hoc analyses were conducted to evaluate whether the relationships between health behaviors and using the Internet for DWPA held for minority groups (ie, non-Hispanic blacks and Hispanics). For the most part, the same positive pattern between health behaviors and use of the Internet for DWPA was observed among Hispanics and non-Hispanic blacks. However, among Hispanics, use of the Internet for DWPA was associated with lower exercise ($\beta=-.05$, $B=-.20$, 95% CI -0.35 to -0.04, $P=.03$); in 2011, Hispanics who used the Internet for DWPA reported 2.49 days of exercise per week. Thus, Internet sites that specifically target Hispanic populations should be mindful of the somewhat lower exercise adherence levels of Hispanics seeking information about DWPA from the Internet.

To test whether the relationships between using the Internet for DWPA and health behaviors changed from 2007 to 2011, we created interaction terms between use of the Internet for DWPA and survey year, and added the interaction terms to the regressions described above. The interaction term was not associated with vegetable intake ($\beta=.001$, $B=.004$, 95% CI -0.08 to 0.09, $P=.83$) and was negatively associated with fruit intake ($\beta=-.07$, $B=-.25$, 95% CI -0.32 to -0.19, $P=.004$) and with moderate exercise ($\beta=-.01$, $B=-.05$, 95% CI -0.10 to 0.00, $P=.05$). Thus, the relationship between using the Internet for DWPA and eating vegetables was similar in 2011 to what it was in 2007, but that the relationships between using the Internet for DWPA and eating fruit or exercising were weaker in 2011 than in 2007.

Discussion

Principal Findings

The current study was the first to examine use of Internet-based DWPA programs in a nationally representative sample of the United States. In general, US adults who were younger, more educated, married, with higher BMIs, and nonwhite were more likely to use the Internet to help with DWPA. The findings for education and BMI are in line with findings from RCTs and studies on commercially available or open-access programs [15,16,24]. While the survey showed that younger people were more likely to use the Internet for DWPA than older people, the mean age of users was 45 (compared to nonusers' mean age of 51), which is also supportive of previous findings. One surprising finding was that women were no more likely than men to use the Internet for DWPA, in contrast to the large percentages of women who enroll in clinical trials and commercially available/open-access websites [14-16,25]. However, changes in demographic characteristics of users changed significantly between 2007 and 2011, with more females, younger adults, and less educated individuals representing a larger proportion of more recent users. Thus, while this gender profile may have changed over time, our results suggest that men use the Internet for DWPA to a greater extent than previous research suggests. Another surprising finding was that non-Hispanic black and Hispanic Internet users were more likely to have used the Internet for DWPA than non-Hispanic whites. The changes over time indicate an increasing proportion of Hispanic users but a decreasing proportion of non-Hispanic black users. Importantly, these findings indicate that the Internet may serve as a useful platform to help address health disparity gaps traditionally found among minority groups. At a broad level, the changes in age, gender, education level, marital status, and race of the typical user point to the dynamic nature of user characteristics, and those who develop and evaluate Internet programs for DWPA should be aware of changes in usership.

A trend indicated that a greater proportion of Internet users used the Internet for DWPA in 2011 than in 2007. This finding provides support for the increasing viability of the Internet as a platform that can reach large numbers of geographically dispersed people. As more people turn to the Internet for help with DWPA, developing and evaluating quality, evidenced-based online programs is of utmost importance.

Importantly, even when controlling for gender, BMI, age, and level of education, people who use the Internet to help with DWPA report greater fruit and vegetable intake and more physical activity than those who do not use the Internet for DWPA. There are a number of possible explanations for this finding. First, it may reflect pre-existing differences in adherence between those who turn to the Internet for DWPA information compared to those who do not. That is, those who use the Internet for DWPA may have greater interest, intent, or motivation towards DWPA than those who do not seek such information. Second, using the Internet for DWPA may help a person better manage healthy behaviors. However, due the correlational nature of the data, we cannot draw conclusions

about the nature of the relationship. To address these causal questions, future studies should examine the relationship between using the Internet for DWPA and related health outcomes in a longitudinal, controlled trial.

We also found that despite more people using the Internet for DWPA in 2011 than in 2007, use in 2011 was associated with lower adherence to healthy DWPA behaviors than in 2007. One potential reason for this finding could be the higher availability of online programs and websites for DWPA; in 2007, it may have taken higher levels of motivation or intention to seek out a website or program for DWPA, and this higher level of motivation may also correspond to better health behaviors. With the proliferation of Internet sites providing DWPA-related information, future users of such sites may be expected to have lower levels of health-behavior adherence than observed in the past. Of course, another possibility is that DWPA-related Internet sites are becoming less effective in their provision of information such that their use leads to lower adherence, but the correlational nature of our data cannot address this possibility. As previously mentioned, only longitudinal, controlled trials can assess the effectiveness of current DWPA Internet sites in improving DWPA behavior.

The US Department of Agriculture and Department of Health and Human Services recommend that adults consume at least 2 cups of fruit and 2.5 cups of vegetables per day given a standard 2000-calorie diet and 150 minutes of moderate physical activity per week [26]. The average self-reported behavior of people who used the Internet for DWPA in 2011 exceeded the recommendations for both fruit and vegetable consumption but

likely fell below the recommendations for physical activity. Thus, it could be suggested that people who currently use the Internet for DWPA use it as a support for generally healthy dietary habits, but not as a tool for improving levels of physical activity.

Limitations

Although this study was the first to assess user characteristics and associated behaviors of using the Internet for DWPA with a nationally representative sample, it is not without limitations. One limitation was the dichotomous nature of our measure for using the Internet for DWPA; there was no way to ascertain an individual's level of use (eg, yearly, weekly, daily), nor to tell the type or quality of the website or program they used. Second, because the 2007 and 2011 HINTS datasets used different samples, we could not make any longitudinal inferences from the data. Last, all data are self-reported, so our data on DWPA-related behaviors cannot be validated against objective measures.

Conclusion

The current study advances prior research by examining the demographic and health behavior correlates of Internet use for DWPA among a nationally representative US sample. As Internet use grows to near ubiquity in developed countries, our findings highlight the importance of evaluating Web-based programs, which may prove effective in health behavior change. The Internet represents a viable platform for targeting health behavior change across a large and growing audience, and future research should continue to explore this important topic.

Conflicts of Interest

None declared.

References

1. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *CMAJ* 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: [10.1503/cmaj.051351](https://doi.org/10.1503/cmaj.051351)] [Medline: [16534088](https://pubmed.ncbi.nlm.nih.gov/16534088/)]
2. World Health Organization (WHO). WHO technical report series number 916). 2003. Diet, nutrition and the prevention of chronic diseases (Report of a joint WHO/FAO expert consultation URL: http://whqlibdoc.who.int/trs/WHO_TRS_916.pdf [accessed 2008-07-21] [WebCite Cache ID 5ZUX4f1O]
3. He FJ, Nowson CA, Lucas M, MacGregor GA. Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. *J Hum Hypertens* 2007 Sep;21(9):717-728. [doi: [10.1038/sj.jhh.1002212](https://doi.org/10.1038/sj.jhh.1002212)] [Medline: [17443205](https://pubmed.ncbi.nlm.nih.gov/17443205/)]
4. Sofi F, Capalbo A, Cesari F, Abbate R, Gensini GF. Physical activity during leisure time and primary prevention of coronary heart disease: an updated meta-analysis of cohort studies. *Eur J Cardiovasc Prev Rehabil* 2008 Jun;15(3):247-257. [doi: [10.1097/HJR.0b013e3282f232ac](https://doi.org/10.1097/HJR.0b013e3282f232ac)] [Medline: [18525378](https://pubmed.ncbi.nlm.nih.gov/18525378/)]
5. O'Connell JB. The economic burden of heart failure. *Clin Cardiol* 2000 Mar;23(3 Suppl):III6-III10. [Medline: [10754775](https://pubmed.ncbi.nlm.nih.gov/10754775/)]
6. Dall TM, Zhang Y, Chen YJ, Quick WW, Yang WG, Fogli J. The economic burden of diabetes. *Health Aff (Millwood)* 2010 Feb;29(2):297-303 [FREE Full text] [doi: [10.1377/hlthaff.2009.0155](https://doi.org/10.1377/hlthaff.2009.0155)] [Medline: [20075080](https://pubmed.ncbi.nlm.nih.gov/20075080/)]
7. Franz MJ, VanWormer JJ, Crain AL, Boucher JL, Histon T, Caplan W, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. *J Am Diet Assoc* 2007 Oct;107(10):1755-1767. [doi: [10.1016/j.jada.2007.07.017](https://doi.org/10.1016/j.jada.2007.07.017)] [Medline: [17904936](https://pubmed.ncbi.nlm.nih.gov/17904936/)]
8. Sacks FM, Bray GA, Carey VJ, Smith SR, Ryan DH, Anton SD, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* 2009 Feb 26;360(9):859-873 [FREE Full text] [doi: [10.1056/NEJMoa0804748](https://doi.org/10.1056/NEJMoa0804748)] [Medline: [19246357](https://pubmed.ncbi.nlm.nih.gov/19246357/)]

9. 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/)]
10. Womble LG, Wadden TA, McGuckin BG, Sargent SL, Rothman RA, Krauthamer-Ewing ES. A randomized controlled trial of a commercial internet weight loss program. *Obes Res* 2004 Jun;12(6):1011-1018. [doi: [10.1038/oby.2004.124](https://doi.org/10.1038/oby.2004.124)] [Medline: [15229342](https://pubmed.ncbi.nlm.nih.gov/15229342/)]
11. Saperstein SL, Atkinson NL, Gold RS. The impact of Internet use for weight loss. *Obes Rev* 2007 Sep;8(5):459-465. [doi: [10.1111/j.1467-789X.2007.00374.x](https://doi.org/10.1111/j.1467-789X.2007.00374.x)] [Medline: [17716303](https://pubmed.ncbi.nlm.nih.gov/17716303/)]
12. Kreuter MW, Strecher VJ, Glassman B. One size does not fit all: the case for tailoring print materials. *Ann Behav Med* 1999;21(4):276-283. [Medline: [10721433](https://pubmed.ncbi.nlm.nih.gov/10721433/)]
13. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. *J Med Internet Res* 2011;13(2):e32 [FREE Full text] [doi: [10.2196/jmir.1624](https://doi.org/10.2196/jmir.1624)] [Medline: [21493191](https://pubmed.ncbi.nlm.nih.gov/21493191/)]
14. Wanner M, Martin-Diener E, Braun-Fahrländer C, Bauer G, Martin BW. Effectiveness of active-online, an individually tailored physical activity intervention, in a real-life setting: randomized controlled trial. *J Med Internet Res* 2009;11(3):e23 [FREE Full text] [doi: [10.2196/jmir.1179](https://doi.org/10.2196/jmir.1179)] [Medline: [19666456](https://pubmed.ncbi.nlm.nih.gov/19666456/)]
15. Anderson-Bill ES, Winett RA, Wojcik JR. Social cognitive determinants of nutrition and physical activity among web-health users enrolling in an online intervention: the influence of social support, self-efficacy, outcome expectations, and self-regulation. *J Med Internet Res* 2011;13(1):e28 [FREE Full text] [doi: [10.2196/jmir.1551](https://doi.org/10.2196/jmir.1551)] [Medline: [21441100](https://pubmed.ncbi.nlm.nih.gov/21441100/)]
16. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. *J Med Internet Res* 2010;12(4):e69 [FREE Full text] [doi: [10.2196/jmir.1640](https://doi.org/10.2196/jmir.1640)] [Medline: [21156470](https://pubmed.ncbi.nlm.nih.gov/21156470/)]
17. U.S. Department of Commerce: National Telecommunications and Information Administration. NTIA research preview. 2011. Digital nation: expanding internet usage URL: http://www.ntia.doc.gov/files/ntia/publications/ntia_internet_use_report_february_2011.pdf [accessed 2013-07-06] [WebCite Cache ID 6HuFzy1vj]
18. Pew Research Center. Pew Internet & American Life Project Surveys, March-April. 2000. Internet adoption, 1995 - 2012 URL: [http://www.pewinternet.org/trend-data-\(adults\)/internet-adoption.aspx](http://www.pewinternet.org/trend-data-(adults)/internet-adoption.aspx) [accessed 2013-02-20] [WebCite Cache ID 6EZwIwPC7]
19. National Cancer Institute. 2012. Health Information National Trends Survey (HINTS) URL: <http://hints.cancer.gov/instrument.aspx> [accessed 2013-02-20] [WebCite Cache ID 6EZwmGSCC]
20. Westat. Rockville, MD. Health Information National Trends Survey (HINTS) 2007 final report URL: <http://hints.cancer.gov/docs/HINTS2007FinalReport.pdf> [accessed 2010-10-08] [WebCite Cache ID 5tKiTZDc5]
21. Westat. Cycle 1 methodology report. Rockville, MD Health Information National Trends Survey 4 (HINTS 4) URL: http://hints.cancer.gov/docs/HINTS4_Cycle1_Methods_Report_revised_Jun2012.pdf [accessed 2013-04-29] [WebCite Cache ID 6GFN0KUDe]
22. Rizzo L, Moser RP, Waldron W, Wang Z, Davis WW. NIH Publication No 08-6435. Analytic methods to examine changes across years using HINTS 2003 and 2005 data URL: http://hints.cancer.gov/docs/HINTS_Data_Users_Handbook-2008.pdf [accessed 2013-04-30] [WebCite Cache ID 6GGth00qr]
23. Tjur T. Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination. *The American Statistician* 2009 Nov;63(4):366-372. [doi: [10.1198/tast.2009.08210](https://doi.org/10.1198/tast.2009.08210)]
24. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. *J Med Internet Res* 2010;12(1):e3 [FREE Full text] [doi: [10.2196/jmir.1361](https://doi.org/10.2196/jmir.1361)] [Medline: [20147006](https://pubmed.ncbi.nlm.nih.gov/20147006/)]
25. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum Internet weight loss program. *J Med Internet Res* 2010;12(1):e9 [FREE Full text] [doi: [10.2196/jmir.1347](https://doi.org/10.2196/jmir.1347)] [Medline: [20350926](https://pubmed.ncbi.nlm.nih.gov/20350926/)]
26. U.S. Department of Agriculture, U.S. Department of Health and Human Services. Dietary guidelines for Americans, 2010. Washington, DC URL: <http://www.cnpp.usda.gov/Publications/DietaryGuidelines/2010/PolicyDoc/PolicyDoc.pdf> [accessed 2011-02-08] [WebCite Cache ID 5wM9GjvZ3]

Abbreviations

DWPA: diet, weight, and physical activity

HINTS: Health Information National Trends Survey

RCT: randomized controlled trial

Edited by G Eysenbach; submitted 12.03.13; peer-reviewed by E Anderson-Bill, J Nunez-Cordoba, P Stumbo, J Duncan; comments to author 12.04.13; revised version received 16.05.13; accepted 11.06.13; published 01.08.13.

Please cite as:

McCully SN, Don BP, Updegraff JA

Using the Internet to Help With Diet, Weight, and Physical Activity: Results From the Health Information National Trends Survey (HINTS)

J Med Internet Res 2013;15(8):e148

URL: <http://www.jmir.org/2013/8/e148/>

doi: [10.2196/jmir.2612](https://doi.org/10.2196/jmir.2612)

PMID: [23906945](https://pubmed.ncbi.nlm.nih.gov/23906945/)

©Scout N McCully, Brian P Don, John A Updegraff. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.08.2013. 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

Consumers' Perceptions of Patient-Accessible Electronic Medical Records

Christina Zarcadoolas¹, PhD; Wendy L Vaughn¹, MPH; Sara J Czaja², PhD; Joslyn Levy³, BSN, MPH; Maxine L Rockoff⁴, PhD

¹CUNY School of Public Health at Hunter College, New York, NY, United States

²Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL, United States

³Joslyn Levy & Associates, LLC, New York, NY, United States

⁴Department of Biomedical Informatics, Columbia University, New York, NY, United States

Corresponding Author:

Christina Zarcadoolas, PhD

CUNY School of Public Health at Hunter College

2180 Third Avenue #547

New York, NY, 10035

United States

Phone: 1 212 396 7745

Fax: 1 212 396 7638

Email: czarcado@hunter.cuny.edu

Abstract

Background: Electronic health information (eHealth) tools for patients, including patient-accessible electronic medical records (patient portals), are proliferating in health care delivery systems nationally. However, there has been very limited study of the perceived utility and functionality of portals, as well as limited assessment of these systems by vulnerable (low education level, racial/ethnic minority) consumers.

Objective: The objective of the study was to identify vulnerable consumers' response to patient portals, their perceived utility and value, as well as their reactions to specific portal functions.

Methods: This qualitative study used 4 focus groups with 28 low education level, English-speaking consumers in June and July 2010, in New York City.

Results: Participants included 10 males and 18 females, ranging in age from 21-63 years; 19 non-Hispanic black, 7 Hispanic, 1 non-Hispanic White and 1 Other. None of the participants had higher than a high school level education, and 13 had less than a high school education. All participants had experience with computers and 26 used the Internet. Major themes were enhanced consumer engagement/patient empowerment, extending the doctor's visit/enhancing communication with health care providers, literacy and health literacy factors, improved prevention and health maintenance, and privacy and security concerns. Consumers were also asked to comment on a number of key portal features. Consumers were most positive about features that increased convenience, such as making appointments and refilling prescriptions. Consumers raised concerns about a number of potential barriers to usage, such as complex language, complex visual layouts, and poor usability features.

Conclusions: Most consumers were enthusiastic about patient portals and perceived that they had great utility and value. Study findings suggest that for patient portals to be effective for all consumers, portals must be designed to be easy to read, visually engaging, and have user-friendly navigation.

(*J Med Internet Res* 2013;15(8):e168) doi:[10.2196/jmir.2507](https://doi.org/10.2196/jmir.2507)

KEYWORDS

electronic medical record; EMR; patient portal; usability; health literacy

Introduction

Background

A number of social and economic factors, such as rising health care costs, a trend towards home health care, as well as shortages of health care workers, have encouraged consumers to increasingly assume a more active role in the management of their own health. Concurrently, advisory bodies, such as the Institute of Medicine, and government agencies have promoted the ability of health information technology to not only enhance the patient-centeredness of health care, but to improve the quality and efficiency of health care delivery overall [1].

One proliferating electronic health (eHealth) tool is the patient-accessible electronic medical record or patient portal. A patient portal, as defined by HealthIT.gov, is an Internet application that allows patients to access their electronic health records and communicate with their health care providers [2]. In this paper, we use patient portal to refer to a secure system tethered to a provider's electronic medical record. The patient portal typically offers patients the ability to communicate with providers, manage medications, schedule appointments, review lab results, access medical history data, as well as provide links to obtain patient education or health information from other online sources [3]. Patient portals allow consumers to take greater control of their health information by changing traditional top-down (doctor to patient) methods of health communication [4] and improving satisfaction with provider communication and overall care [5]. While research is limited, some studies suggest benefits such as greater engagement in health care with online access to personal health information [6], improved rates of screening [7], improved appointment adherence [8], and greater sense of confidence and empowerment, increased knowledge about health, and improved health behaviors [9,10].

The Health Information Technology for Economic and Clinical Health (HITECH) Act has promoted the adoption of health information technology by incentivizing the meaningful use of electronic medical records. Tying incentive payments from Medicare and Medicaid to meeting meaningful use objectives, newly released requirements stipulate that patients must be given online access to their health record with the ability to access, print, share, or download their health information. In addition, providers will need to ensure that at least 5% of patients actively use this technology [11].

The study is motivated by no less than two factors. First, consumers are increasingly interested in accessing their personal health information online [12-16]. The Markle Survey on Health in a Networked Life reports that roughly 70% of the public and 65% of doctors believe that patients should be able to download and keep their own health information [17]. Walker and colleagues (2009) conducted focus groups with frequent Internet users and found that participants not only wanted electronic access to their medical records, but customized health information and advice as well [18].

Second, the literature has documented the vulnerability of certain population groups to disparities in health outcomes [19,20] and

health care quality [21,22]. Vulnerable groups are most often described as racial and ethnic minorities, poor, under-educated, immigrants, and those lacking English proficiency [23]. Many vulnerable populations are also likely to be low literate and/or low health literate [24,25]. Strong evidence exists linking low literacy and low health literacy with poorer health behaviors, disease management skills, and health outcomes [26-31]. This is particularly concerning considering most health information created for the general public is written at or above the 10th grade level [32], with health websites often at an even higher level [33], even though more than half of the adults in the United States read at an 8th grade level and lower, and some vulnerable populations read at 5th grade level or lower [34].

Vulnerable populations continue to receive poor quality health care as well as face more barriers to receiving care than more advantaged groups, despite continuing efforts to reduce such disparities [22]. Some have suggested that use of eHealth, by providing increased access to health information and support, may help to ameliorate disparities in vulnerable or disadvantaged groups [23,35]. However, there is conflicting data regarding these consumers' perceptions and the use of patient portals and other electronic health information.

Studies with medically underserved and vulnerable patients [36,37] found strong interest in accessing online health records, and research also shows that some vulnerable populations, such as racial and ethnic minorities, have similar or even greater interest in accessing health information online than national samples [38]. Other studies show that racial and ethnic minorities have lower rates of enrollment [39-41], logging on [42,43], likelihood of receiving an access code, and regularly using a patient portal [44] than non-minorities. However, once enrolled, studies found no disparities in use by race/ethnicity [39,45] or less of an association with race/ethnicity than with portal adoption [41]. These findings indicate that the way patient portals are designed and presented to consumers may influence how portals are perceived, valued, and ultimately utilized.

Objective

Thus, our goal in this qualitative study was to identify what perceived utility and value vulnerable health consumers attach to the concept of patient portals and to core features and functions of these portals. We utilized focus groups targeting lower education level, minority residents of New York City to explore this question.

Methods

Participants

A total of 28 individuals, 10 male and 18 female, 21-63 years of age, with a high school education or less, participated in 4 focus groups held June-July 2010. Each of the 4 focus groups consisted of a convenience sample of 6-8 participants along with a moderator and a note taker. The groups were conducted using a semistructured format. Prior to commencing the focus groups, participants completed a technology experience questionnaire adapted from Czaja et al [46,47]. Each group lasted approximately 90 minutes. All focus groups were audiotaped. At the end of each group, participants were paid

US \$40.00 and given a US \$4.50 round-trip NYC MetroCard for their time and travel. All study protocols and materials were approved by the Mount Sinai School of Medicine Institutional Review Board.

Sample

Individuals were eligible to participate if they met the following criteria: New York City resident, aged 21-75 years, able to read and speak English, and no higher than a high school education/GED.

Recruitment

Recruitment was conducted at 3 sites in New York City: Mount Sinai Medical Center, East Harlem; Queens Library, Long Island City, Queens; and CAMBA, Flatbush, Brooklyn. These organizations and locations were chosen for their access to diverse population groups in different urban neighborhoods in New York City.

Each site is located in a health service area designated as Medically Underserved and a Primary Care Health Professional Shortage area by the United States Department of Health and Human Services Health Resources and Services Administration (HRSA). According to HRSA, Medically Underserved Areas have “too few primary care providers, high infant mortality, high poverty and/or high elderly population” [48].

A description of the focus group neighborhood, site, and participants can be found in [Table 1](#) [49-53].

We specifically recruited in Queens for 2 age groups: one 18-49 and the other, 50 and older. We chose this location because we believed they had the largest and most age-diverse pool of candidates meeting our criteria. Although there are a number of different age ranges within which researchers and others operationalize “younger” and “older” adults, we follow other studies and surveys, which have defined “older” adults as those 50 years and older [54-56].

Using IRB (Institutional Review Board) approved recruitment and advertising materials, staff members of these programs assisted with recruitment of participants by distributing flyers

about the study, including the inclusion criteria, to their participants and patrons. Interested individuals either signed a sheet requesting researchers contact them to be screened for the study or called researchers directly. Potential participants were screened by research staff over the phone to ensure they were eligible for the study. Thirty-two individuals met eligibility criteria, agreed to participate, and were scheduled to attend a focus group. Twenty-eight participants actually attended a group. All participants provided informed consent according to IRB requirements.

Procedure for Focus Groups

It is often the case in focus group research that participants are somewhat familiar with the topic of discussion. However, in this case, most participants were unfamiliar with the concept of patient portals and their specific capabilities. Therefore, the moderator provided a scripted introduction to patient portals including demonstrations using readily available informational videos as prompts. The videos were produced by a large health system, a managed care company, a multispecialty health care provider, and a federally qualified health center network—all of whom have been forerunners in the adoption of electronic medical records for patients. We selected clips from these organizations because they represented a broad range of possible entities from which patients may receive health care. The videos were created by these entities for their patients, so what participants saw was what a real world patient might encounter when being introduced to a portal for the first time.

The moderator’s guide focused on exploring the following four core functions of portals: scheduling appointments, managing medications, proxy functions, and reviewing lab results. These four functions were chosen because they are universally present in most patient portals. Participants were asked questions such as what they thought about the idea of portals in general and each of the four functions and how likely would they be to use the feature. Participants were also encouraged to ask questions and offer critiques of portal content and capabilities, as well as respond to a scenario that asked how a patient portal might be of use in keeping them healthy.

Table 1. Focus group sites and participants.

Neighborhood	Snapshot	Focus Group Site	Description	Participants
East Harlem, Manhattan	38% of East Harlem residents live below the poverty level and 69% have a high school education or less. The population is 55% Hispanic and 33% black/African American [49].	Mount Sinai Medical Center, Department of Health Education	The Department of Health Education provides comprehensive education and support programs in the community.	6 women
Long Island City, Queens	Long Island City is located in Northwest Queens. 20% of residents in this area live below the poverty level and 55% have a high school education or less. The population in Northwest Queens is 43% White, 28% Hispanic, 15% Asian, and 6% black/African American [50]. The census tract in which the library is located is 44.9% Hispanic and 39.8% black/African American [51].	Queens Library, Long Island City Branch	Queens Library system serves the most ethnically diverse county in the United States and has the largest circulation of any US public library. The Long Island City branch provides a range of services, from adult learning and literacy programs to job readiness and computer training.	(1) 8 adults (2 men, 6 women), all between 18-49 years, and (2) 7 adults (2 men, 5 women), all 50 years and older
Flatbush, Brooklyn	21% of Flatbush residents live below the poverty level and 58% have a high school education or less. The population is 77% black/African American and 9% Hispanic [52]. Over half of Flatbush residents are foreign-born (51%), primarily (80%) from non-Hispanic Caribbean countries [52,53].	CAMBA	CAMBA is a community-based organization in Brooklyn committed to serving low-income, homeless, and immigrant populations, among other vulnerable groups.	7 adults (6 men and 1 woman)

Analysis

All focus groups were written up in quick notes by both the moderator and observer independently within 24 hours. Quick notes record initial impressions as these can extinguish rapidly over time. The analysis team, consisting of the moderator and note taker, established a process for analysis that consisted of developing a coding guide as follows. We used a grounded theory (GT) model to analyze focus group data [57-59]. GT involves constantly comparing the data, coding, and identifying interchangeable indicators to reveal patterns that ultimately lead to categories. Two coders independently and repeatedly listened to the audio, referred to quick notes, and then conferred to discuss and refine emerging topics and themes. This process continued until the coding guide was finalized. The coding guide and accompanying narrative summary was developed to describe the key content of each group and enable a ready comparison of break characteristics, that is, the factors that differentiated one group from another.

All 28 participants had no higher than a high school degree, 15 had a high school diploma or GED, and 13 did not complete high school. Most were of ethnic/racial minority backgrounds, 7 Hispanic, 19 non-Hispanic black, 1 non-Hispanic white, and 1 Other, and all were of low economic means, including 8 with reported household incomes below US \$20,000 (Table 2). Although more than half of participants refused or said that they did not know their annual household income, it is acknowledged that research participants may not provide income information because they are unsure of the answer or feel it is too private to share [60]. However, research has also shown that those who refuse to provide income information are more likely to not be working, have less education, and live in a low socioeconomic neighborhood than those who report income [61]. Given the neighborhood characteristics of our focus group locations, we are confident that our participants are representative of the vulnerable population we intended to reach.

Table 2. Participant demographics (N=28).

	East Harlem (n=6)	Brooklyn (n=7)	Queens (n=8)	Queens older adults (n=7)	Total (n=28)
Age in years, mean (SD)	26.7 (5.7)	37.1 (5.6)	38.0 (9.5)	56.6 (4.2)	40.0 (12.4)
Gender, n					
Male	0	6	2	2	10
Female	6	1	6	5	18
Ethnicity/Race, n					
Hispanic	5	0	2	0	7
Non-Hispanic white	0	0	0	1	1
Non-Hispanic black	1	7	5	6	19
Other	0	0	1	0	1
Education, n					
Less than HS	3	2	5	3	13
HS graduate/GED	3	5	3	4	15
Income, n					
<\$20,000	1	1	3	3	8
\$20,000+	3	0	1	1	5
Don't know / refuse	2	6	4	3	15
Occupational status, n					
Work full-time	3	3	0	0	6
Work part-time	1	0	0	0	1
Student	0	0	3	0	3
Homemaker	0	0	1	0	1
Retired	0	0	0	3	3
Seeking employment	2	4	4	2	12
Other	0	0	0	2	2
Primary language, n					
Yes (English)	6	0	8	7	21
No (Other)	0	7	0	0	7
General health, n					
Excellent	0	3	1	1	5
Very good	1	0	2	1	4
Good	3	3	3	3	12
Fair	2	0	2	1	5
Poor	0	1	0	1	2

Results

Technology in People's Daily Lives

First, we describe participants' current technology and eHealth use, record-keeping behaviors, and health information seeking behaviors. Then, we discuss participants' responses to selected features of patient portals. Finally, we present themes arising from participants' attitudes and perceptions of patient portals.

The majority of participants currently used communication technology such as computers, the Internet, and mobile devices. All reported that they were currently using or had used a computer in the past and only 2 participants reported having no Internet experience (Table 3). All participants had either a basic cell phone or smart phone (some participants used both) (Table 4).

Most participants used a computer at home, but many also accessed computers in a variety of other places, such as at a public library, adult learning center, friend or relative's house,

work, or community center (Table 5). Many used technology to perform tasks of everyday life such as sending email, accessing social media, playing games, and searching for information online (Table 6). Most participants were not only technology users but were interested in learning more ways to stay connected through technology.

The East Harlem group was the most facile and “on the grid” of all of our participants. All 6 women regularly reported using a computer and the Internet in their daily lives. Facebook, email, pictures, Myspace, shopping, Twitter, school websites to check their children’s progress, planning travel/trips, or finding their way around town were among the things that they routinely do online. The Queens older adults group reported that they were learning how to access new features on their mobile devices or computers, with several expressing a desire to learn how to “text” and others regularly using social media. Members of the Brooklyn group were somewhat less likely to have used some forms of technology (the only group with some members reporting no Internet experience—see Table 2). However, they were more anxious to learn about and embrace technology than the Queens older adults, perhaps because they saw mastery of technology as a key part of their acculturation and social and economic success.

Previous use of technology for health-related issues was limited across all groups. Only 10 of 26 Internet-using participants reported ever having searched for health information online (Table 6), in contrast to national statistics showing 80% of all online adults have used the Internet to search for health information [62]. However, there were a few participants who reported not only searching for health information, but accessing their own personal health information electronically. One

participant (Queens 21-49) reported having access to a fully functioning patient portal, noting that she had a “health scare”, for which she regularly accessed a portal to help her “stay on top of my medical records”. Another participant in the Queens older adults group reported carrying his medical information on a jump drive attached to his keychain.

Most participants across groups recognized benefits in using technology and eHealth tools to manage their health care, although a few expressed concerns associated with these tools. Younger participants, such as those in the East Harlem group, were very comfortable with technology and saw the use of eHealth tools as an extension of what they were already doing in their everyday lives. They also expressed the most interest and excitement around features that would make their lives more “hassle free”, such as using the portal to make appointments and get copies of medical and immunization records for their children.

Likewise, participants accustomed to being online and using social media transferred their expectations for user-friendly formats and tools to the patient portals we showed them. Participants expressed a desire for features such as mouseovers/clicks for just-in-time information. They faulted a site for not having such common features. For some, having a website interface that resembled commonly used Internet sites affected how they felt toward the patient portals we were displaying. One participant summed this up by saying that if someone were trying to sell him on one of the systems we demonstrated, as an “average person”, he would pick the one that put “everything out front...it’s like working on Facebook or Hotmail, everything is in plain sight and I can deal with that and most people like that”.

Table 3. Use of computers and the Internet (N=28).

	East Harlem (n=6)	Brooklyn (n=7)	Queens (n=8)	Queens older Adults (n=7)	Total (n=28)
Length of time using a computer, n					
> 5 years	6	2	6	4	18
>1 year, but <5 years	0	1	0	1	3
Between 6 months and 1 year	0	2	2	1	4
<6 months	0	2	0	1	3
No computer experience	0	0	0	0	0
Experience with the Internet, n					
Internet experience	6	5	8	7	26
No Internet experience	0	2	0	0	2

Table 4. Use of technology (N=28) of participants who use “frequently”.^a

Technology	n (%)
Basic cell phone	26 (93)
Recording and playback device	20 (71)
ATM	17 (61)
TV set top box	16 (57)
Home telephone	13 (46)
Smart phone with Internet access	10 (36)
Computer/Video games	10 (36)
Digital photography ^b	10 (37)
MP3/iPod music player ^b	10 (37)
In-car navigation system	4 (14)
Automated movie ticket purchase kiosk	3 (11)
Fitness devices	3 (11)

^aPercentages sum to more than 100% because participants could select multiple responses.

^bOne participant did not respond (n=27).

Table 5. Location of computer use (n=27^a).^b

Location	n (%)
Home	18 (67)
Public library	10 (37)
Adult learning center	9 (33)
Friend or relative's house	9 (33)
Work	7 (26)
Community center	3 (11)
Other	3 (11)

^aOne participant did not respond.

^bPercentages sum to more than 100% because participants could select multiple responses.

Table 6. Activities on the Internet % (n=26) of participants who use “frequently”.^a

Activities on the Internet	n (%)
Email	17 (65)
Read the news online	13 (50)
Instant messaging ^b	12 (48)
Post resume or search for employment	12 (46)
Social media ^b	11 (44)
Find information about community events or religious services	10 (39)
Search for health information about an illness, or order medications or other health products online	10 (39)
Make reservations, search for maps, or get travel information online	10 (39)
Search for educational courses or materials, use instructional software, or participate in online degree training programs	9 (35)
Shop for clothes or other items, search for product information online	9 (35)
Download government forms or find information about benefits or programs ^b	8 (32)
Buy tickets, find information about shows, events or hobbies	5 (19)
Online banking and/or bill paying	4 (15)

^aPercentages sum to more than 100% because participants could select multiple responses.

^bOne participant did not respond (n=25).

Record-Keeping Behaviors

We asked people about what kind of record keeper they would say they were, in order to compare this information to their level of enthusiasm about an electronic medical record. No strong connection was found. The groups differed in self-description of their general record-keeping behavior. Most participants, with the exception of the East Harlem group, asserted that they kept records in their daily lives. Some participants, especially those in the Queens older adults group, noted that they kept “paper records”, for things like tax purposes and appointments. Several participants across groups noted that they kept records of their health information: “I keep medical records, my copies of labs and test results from doctors”. Another participant, as noted earlier, kept his medical information electronically on a jump drive.

Sources of Health Information

When asked about sources of health information, participants reported using a mix of offline and online sources. Offline, participants reported going to the library, pharmacies, and health fairs, or, as one participant described “just talk[ing] to people who have a condition or a health care provider to get more information”. Of the online sources, WebMD and Google searches were most frequently mentioned as ways to obtain health information. However, although some participants commonly used the Internet to search for health information, several commented on the negative aspects of online health information. One participant described his experience with WebMD: “That thing is so hard to understand because they give you a schematic of a human body and you have to point to it but...it doesn't really break it down where the average person can understand...some of the translation is in doctor terms, not in layman terms, so the average person that's looking at it gets lost”. Others noted that Google searches result in “getting too

much information” or finding searches “a little frustrating ’cause it's not always exactly what you're looking for, a specific answer to your problem”.

Initial Participant Response to the Idea of Patient Portals

After being given a short overview of patient portals, participants were asked to respond to the concept of a portal by rating the importance of access to a portal on a scale of 1-5 (with 1-Not Important to 5-Very Important). Most people across groups rated access to a portal either Very Important or Important to them, with the exception of the Queens older adults group. Participants in the East Harlem group embraced the concept from the outset, with all members rating potential access to a portal as “5-Very Important”, identifying benefits in the use of technology and patient portals to manage health care for both their children and themselves. These women envisioned the benefit in having remote access to test results, immunization records for children, and reminders for upcoming appointments (“I think it is great”). The majority of participants in the Brooklyn and Queens (21-49) groups rated patient portals highly, either a 4 or a 5. Members of the Queens older adults group had the least initial enthusiasm towards portals. Only 1 participant in this group rated access to a patient portal as “5-Very Important”, with 4 participants rating access to patient portals as “1-Not Important”.

Next, the groups were shown short videos demonstrating appointment setting, health proxy functions, medication management, and lab test results from a selection of patient portals.

Response to Patient Portals: “Housekeeping” Tasks

Appointments

Almost all participants were impressed with the convenience and control portals provided for making appointments with their health care provider. Many reported frustration with the interactive voice response (IVR) systems they encountered when calling their providers and appreciated the convenience and control offered by a portal—“This seems to give more control over when you would like to make an appointment”. However, some participants found the “drop-down menu” a barrier to use. They were not sure that the appropriate options would be listed in such a menu or that the available choices would allow them to fully describe the nature or urgency of their complaint—“I would just be concerned about if it’s verified enough what you could put, the reasoning, that’s the only thing...when you speak with [a staff person] they might help you be more specific”. Thus, although participants appreciated the utility of the appointment-making feature, they still wanted to know that they had the option of calling their provider’s office directly.

Health Proxy Function

Many participants saw value in having proxy access to the electronic records of their children or their parents. Functions that would enable them to monitor their relative’s health, manage the medications their relative was taking, and understand more about how a relative’s chronic conditions should be managed on a regular basis were seen as especially valuable. In reference to caring for her elderly parent, one participant commented, “I’m dealing with my father who’s been having a stroke. I’m second health proxy and there’s three of us...[it is] better for communication”. Mothers in the East Harlem group felt the proxy function would increase convenience, especially if the portal allowed them to access health and vaccination records on the spot. Said one mother, “I have two kids and just with school and everything, they need physicals, immunizations, everything, if I can access that online, it would just cut out so much of the back and forth [participant in background, ‘hassle’], to actually have to go to your doctor’s office and wait for an hour just to get their records”.

Some consumers in the Queens 21-49 and Queens older adults groups were more circumspect about the possibility of an adult child or other family member gaining access: “Suppose you, as a parent, don’t want your child to know what is going on with you?” Continued discussion, however, assuaged some of these concerns, for example, understanding that the proxy is a voluntary act and that patients did not have to allow anyone, including their grown children, electronic access to their record.

When participants pondered the issue of parental access to an adolescent’s online health records, participants across all groups, except for the Brooklyn group, found privacy and confidentiality issues complicated to resolve. One participant believed that a parent having access to their teenager’s patient portal could inhibit the teen from talking to his or her doctor: “It might just make them more hesitant to disclose things to their doctor if they know that their parents can view this... it might just make them not want to say things that they don’t want their parents to know”. Another participant noted that even in current medical

practice, “with teenagers of a certain age...they do have something where the parents are not allow[ed] to see certain things”. Thus, many participants believed a portal’s proxy function could be useful but acknowledged that rules concerning appropriate access needed further consideration.

Response to Patient Portals: Health Management Tasks

Medication Management

Use of a portal for medication management tasks, such as requesting medication refills, was not as popular a feature for participants. For some, the importance of the feature was not immediately apparent because, in the words of one participant, “I don’t get no medicine”. Many participants seemed more comfortable calling to get a prescription refilled or going directly to the pharmacy, noting that you may not always have access to a computer. However, the one participant who had access to a patient portal reported that she found this feature to be a great benefit, “With my doctor, if I need medication it goes straight to the pharmacy. It’s emailed right away. So it is excellent. Without having to wait for a paper prescription...By the time you get home you could pick your medicine up, so it’s very convenient”. Even those participants who did not feel that the medication management feature would be a benefit for them felt that it would be good for the “disabled and elderly”.

Lab and Test Results

Access to lab and test results is a central function of most patient portals. Most participants said being able to see their test results was a very important feature for them. All groups viewed the following example of a lab test result from a patient portal (typical of test result formats we have reviewed) (Figure 1).

“Where is the standard?...The 0-7%, what exactly does that mean?” and “Is [that a] normal level for my blood?...I don’t know”. Participants wanted this page to be easier to use: “[be] more specific, like kind of in layman terms...Is that a good result? Is that a bad result?” or state clearly, “For a healthy person, it should be...you’re either above the window or below it”.

When participants were shown a test result that included a short explanatory note from the doctor (Figures 2 and 3), they felt they better understood the test result numbers and whether they should be concerned or not, noting that, “It’s written. It’s clear. You can look at it and understand it. Sometimes the doctor’s visit is so fast you might not even get as much as that [in your visit]”. Another said, “A lot of times...the doctor explain[s] things, but he don’t explain things...99% of the time when people go to the doctor they want to get in and out...with this here, I truly feel because they broke [it] down. Okay, this is high but...you have a reminder, you have something...that’s there that you could look at.”

For some consumers, however, the accompanying doctor’s explanation did not go far enough. They wanted the test results page to explain the doctor’s recommendations, for example, that the cholesterol test results were not high enough to require medication and to change their diet, and a further explanation of the different test components and the risks associated with certain values in a way they could understand—the why behind

the recommendation. Looking at the sample results for a cholesterol test, one participant commented, “See this, the numbers...[referring to the doctor’s note] he just telling me that

other part, making me feel comfortable, but I want to know what’s that 210 and...what’s the difference between the LDL and the HDL...break it down for me”.

Figure 1. Sample hemoglobin test result from patient portal.

Component Results		
Component	Your Value	Standard Range
HEMOGLOBIN A1C	8	0 - 7 %

Figure 2. Sample cholesterol test result.

Component Results		
Component	Your Value	Standard Range
CHOLESTEROL	210	100 - 199 mg/dL
LDL CHOLESTEROL	141	0 - 99 mg/dL
HDL CHOLESTEROL	43	40 - 59 mg/dL
TRIGLYCERIDES	183	- <150 mg/dL

Figure 3. Provider's note accompanying cholesterol test result.

Hi Michelle Zztest,

Your cholesterol result is high but not high enough that I would start a medication at this time.

I would recommend that you try to eat a healthier diet and begin exercising.

- Try to eat less fried foods, butter, oils, cheeses and fast food.
- If you are going to use oil, switch to olive oil which is healthier than regular vegetable oil.
- Also try to eat more fruits and vegetables.
- Exercise will help raise your HDL (good cholesterol).
- You can also try using a dietary supplement called omega 3 fish oils which could also help improve your cholesterol

We should recheck your cholesterol in 3-6 months.

Regards,

Key Themes

Consumer/Patient Empowerment—“Information is Power”

As reported above, most participants, with the exception of the Queens older adults group, were very interested in having access to their personal health information from the start. Many individuals expressed what can only be described as disbelief

that the difficulties they historically have had accessing their medical records would be remedied. Participants recalled encountering cumbersome and daunting bureaucracies when they wanted a copy of their own medical records or that of a care recipient, such as a child or elderly parent (“Gotta go sign the paper, and you gotta wait a week, and then you gotta pick it up”). Others reported having to pay per page for copies of their medical record or were charged a co-pay to get test results.

Extending the Doctor's Visit/Enhancing Communication With Health Care Providers

Participants almost unanimously named the short length of a doctor's visit, the stress of not understanding things fully while at the visit, and the inability to access useful information post visit as characteristic of their health care experience. Patient portals were seen as offering at least a partial remedy for this very unsatisfactory situation. Many felt that online records could empower them with increased access to health information, resulting in a greater focus on their health and allowing them to be more proactive about taking care of their health.

The efficiency and portability of electronic medical records was seen as another huge benefit. Many participants said that their doctors used electronic medical records, and they felt that these records enhanced communication with providers: "When it was paper, when you go see a new doctor, they're like asking you everything that you already given, all your demographics...when they can just look it up...and they're not asking you all the same things all the time". Although a few participants described themselves as "old school", preferring in-person doctor-patient encounters without an electronic record ("the whole everything-on-the-computer thing is cool, it's okay, but I'm kind of old school, I like the [personal]") or accessing health records in paper format, there was also an acknowledgement that times change and with those changes would come greater acceptance of new technology ("For the next generation, it could be accepted...a lot of things that were thought about in the 60s is now being implemented in 2000"). Several participants, in fact, could foresee the benefits of enhanced communication with health care providers due to the portability of health information via patient portal: "I like it because if you go on vacation, you can get sick. And you have your medical record [and] you can give it to any doctor". Another could see the utility of a portal to enable her to get more out of a doctor's visit with her young children, "When I take my six-year-old to [an] appointment, I usually have my three-year-old there. They're running around. You're trying to listen to what the doctor's telling you, but you're still focused on, you know, your other kid. If you can go home and read everything...you're more knowledgeable".

Literacy and Health Literacy

Participants talked about the need to have patient portal information presented in ways "the average person can understand". For example, after showing participants the sample lab test result page, one participant commented that she would use it "if I could read it and understand it...[but] I can't understand it". Another participant commented, "I just look and see numbers [referring to the sample results page]". The reading and health literacy load (the demands on a reader's understanding of science and other related concepts) of the sample lab test results in the demonstrated patient portal proved challenging and frustrating for participants. Participants found the reading and health literacy load of the content too high in 3 broad categories: (1) medical terms/medical information, eg, drug names, anatomy, chemistry, medical procedures ("There's some people that don't know the difference between good cholesterol and bad cholesterol. See, I didn't know that until I

got into my forties. Always just thought cholesterol. So that's something they would have to explain"), (2) numeracy/numerical information, eg, number calculations, standard range, percentages ("Even if it does say...normal range, you're still up by 1, your value, so that would kind of make some people freak out, like ok, what does that mean?"), and (3) design, navigation, and aiding tools, for which more than one participant suggested a design modification, such as an aiding tool that would allow you to click on the name of the test to get a more detailed explanation ("Could you click on each one and it can give you...a Wikipedia of it...?").

Prevention and Health Maintenance

In the last segment of the focus groups, we asked participants to talk about what role portals might play in a user's health. They commented that using the patient portal would allow people to "stay on top of [their] health", "[focus on] prevention", and "know more". Although most were now positive about portals, they spoke in generalities about how a person's health might improve as a result of using one. Most salient were the positive health impacts of reminders for appointments, annual visits, and screenings. Said one participant, "a lot of people don't take charge of their health because they don't even remember to take care of themselves, and a lot of times they don't even know at what age they should be checking for what things".

Privacy and Security Concerns

While privacy and security were not high priority concerns for most consumers we spoke with, in 3 of the 4 groups (with the exception of the Brooklyn group) there were a small number who voiced concerns related to the privacy and security of online records. An East Harlem participant, who asserted that access to a patient portal was very important, did so with the caveat: "If they are at your doctor's office, you know anybody that has access to it, it's like staff mainly, if it's [a] website and...if there is some way someone can get your password...that's a lot of information for someone else to have. Do you want to risk that?" A participant in the Queens 21-49 group feared that if someone was going through a divorce, an unhappy spouse might gain access and try to "damage the medical record" out of "spite". Several members of the Queens older adults group feared "hacking" and identify theft; a couple of these participants initially felt that concerns over the privacy and security of electronic records would prohibit them from using a patient portal: "It comes down to a security point...the computer is awesome, I mean it's off the chain. You could get on that computer, you could find out about anything, but do you really want your information up there?...I mean it's a question".

Some security and privacy concerns lessened once specific security features associated with patient portals were discussed, and participants identified the potential benefits of patient portals in their daily lives. Participants appeared willing to make tradeoffs, accepting the potential risk of breaches to their personal privacy for the convenience and accessibility of electronic records. Across groups, only one participant maintained throughout that privacy and security concerns would prevent her from using a patient portal, but she acknowledged portals should be an option for those who wanted them.

Response to the Concept of Patient Portals With More Information: Post Demonstration

After participants had the opportunity to view and discuss all of the demonstrations, they were asked to revisit their initial opinions of patient portals. Most participants, who were either indifferent or negative about the value of portals at the outset, raised their opinion of the importance of access to a portal once they had an opportunity to view and discuss common features; this was especially true in the Queens older adults group. Others, awakened to the possibilities, looked to the possible future benefits: “Is it possible, after let’s say, like your relatives, a parent or what have you, pass away, are you able to have access so that doctors can see what diseases and things that run in the family or...predisposition to be exposed to certain things?” A few wanted to know when a patient portal would be set up by their health care providers, so they could begin to access their records electronically.

Discussion

Principal Findings

Online technologies continue to change society and our daily lives in many ways. We see this graphically in the exponential growth in consumer use of the Internet and mobile devices to access health information and services [63-66]. So too, the changes that patient electronic medical records portend are likely to change basic paradigms of patient provider interactions and the historic alignment of access to information.

Our goal in this study was to identify the perceived utility and value vulnerable health consumers attach to the concept of patient portals as well as to core features and functions of these portals: appointment setting, proxy functions, medication management, lab results, and preventive information. Our method was conducting a series of focus groups targeting lower education level, minority residents of New York City.

As expected, most consumers we spoke with were not familiar with a patient portal and only 1 participant out of 28 regularly used a portal. Despite this, most were positive about patient portals and positive perceptions increased over the course of the focus group discussion so that those less positive at the start, became more so as the groups evolved.

Our participants did not indicate that lack of Internet access or privacy when using public computers to access portals posed a potential barrier to portal use, as found elsewhere [36,67]. Nor did our participants appear to have issues resulting from a lack of computer skills or experience, as has also been reported [68]. This may be due the fact that all of our participants had a fair degree of prior experience with technology: all had previously used computers, almost all had experience using the Internet, and a majority had in-home Internet access.

As discussed in the Methods section, participants were introduced to a range of portal features and were asked to reflect on the utility and values of these features. Throughout the sessions, participants were engaged, asked frequent questions, and made extensive comments and critiques of specific features of portal features and functionality. Participants listened

carefully to the descriptions and demonstrations of features, such as making appointments and reviewing lab tests. Then they posed their own questions, for example, “would you receive a confirmation of your appointment?”, “will the portal tell you whether that cholesterol number is ok?”, and, “if not, what you should do about it?” In contrast to perceptions by some consumers that it would be easy to learn how to use electronic medical records [36], participants in these groups identified a number of health literacy and usability barriers to patient portal use, such as complicated medical and numerical information, as well as a lack of aiding tools.

On a broader level, participants easily recognized the ability of electronic records to empower them through increased access to their own records and as a way to get more out of a doctor’s visit. Agarwal and colleagues found that patients who perceived electronic records as empowering had significantly higher intentions to use a portal [69]. In our study, only a few expressed concern that electronic medical records would diminish their relationship with their provider. Instead, most participants focused on the benefits of portal use, especially in light of their increasing frustration with the ever shortening time with the doctor and growing complexity of the health information presented in visits. Research suggests that even if patients were able to understand all of the information related to their visit, only about half of the information would be remembered [70,71], if remembered correctly at all [72]. Thus, introducing patient portals to these consumers allowed them to foresee using a portal as a tool to reduce the burden of remembering everything shared during a doctor’s visit by allowing access and retrieval of visit related information at their convenience. Changing perceptions so that more patients begin to view electronic medical records as an “extension” of the doctor’s visit and as a complementary tool, may ease concerns of those who worry that use of electronic records will supplant or depersonalize provider relationships.

The majority of participants judged their current methods of engaging with their health and “staying on top” of their health as in need of much improvement. Although participants offered up only a few tangible specifics, most perceived patient portals as a good tool for improving their knowledge and engagement in their health care and that of their loved ones.

Limitations

As with all qualitative methods, focus groups are not representative of any larger group of people. It would be important and enlightening to conduct further qualitative study to see if these initial findings regarding interest, empowerment, most appealing functions, barriers to use, and other reactions hold up. Focus group participants did not readily introduce specific ways that their health behaviors might change as a result of using a patient portal. In part, this could be a product of the questions used in the moderator’s guide. We think it is likely that the artificial nature of the portal review in this setting did not provide enough specific and personalized content and context for participants to more substantively reflect on what they, as a real patient with their own health conditions, would most likely do with a patient portal.

Conclusion

The intent and promise of patient portals is that they will help engage people with their health, improve preventive care behaviors, and permit better management of chronic conditions. A handful of recent studies have begun to examine patients' uptake of portals as well as patients' perceptions and assessments of their actual use. Consistent with the views of our study participants, users report finding great value in patient portals [73], a perception shared by a wide range of patients, including those with mental health and substance abuse issues [68] as well as patients with HIV [67]. In general, patients report that portals positively impact communication with providers, and improved knowledge, empowerment, and self-care [73,74].

However, despite the positive potential of portals suggested by these findings, there still remains relatively little exploration of vulnerable patients' experiences with these systems. In order to realize these benefits for all consumers, it is important to better understand the perceived and real barriers and opportunities that vulnerable groups, so disproportionately impacted by health disparities, face in actively using patient portals. This project moves a step beyond surveys and analyses about consumer use of patient portals by employing exploratory focus groups to study the perceptions and opinions of vulnerable, low educational level, and ethnic minority consumers in underserved communities. The findings signal that there is little reason to assume vulnerable populations are not accessing patient portals because they do not see the value of having

electronic access to their personal health information. To the contrary, once consumers were given the opportunity to view and experience portal functions, they became very interested and motivated. Most participants perceived patient portals as a way to finally obtain information that has, until now, been out of their reach; as they often said, "knowledge is power." We hypothesize that this target population, in bringing the experience and perspective that is a manifestation of their historically marginalized status, is, perhaps, even more motivated to embrace this and other emergent technologies.

These focus group findings serve as one baseline representation of public opinion regarding patient portals. Further study of public perception and preferences surrounding patient portals can be used to better guide the introduction of this technology to patients and add argument to the importance of promoting patient portal adoption and use. For example, some have suggested that increasing mobile access may encourage initial adoption and uptake of portals, since minorities are more likely to use these devices for health applications [43]. In addition, our results demonstrate that portal content must be developed that accommodates the needs of lower education level, low literate, and low health literate users. Clearly, it is imperative that patient portals are designed and refined with the input of a diverse body of consumers in order to guard against continued barriers and poor access to information. If not, patient portals will remain an unused or underutilized tool for those who could potentially benefit the most.

Acknowledgments

The project described was supported by Award Number R21CA133487 from the National Cancer Institute. The manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: Committee on Quality of Health Care in America, National Academy Press; 2001.
2. HealthIt.gov. What is a patient portal?. URL: <http://www.healthit.gov/providers-professionals/frequently-asked-questions> [accessed 2013-03-17] [WebCite Cache ID 6FC6gb3Si]
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. Kreps GL, Neuhauser L. New directions in eHealth communication: opportunities and challenges. *Patient Educ Couns* 2010 Mar;78(3):329-336. [doi: [10.1016/j.pec.2010.01.013](https://doi.org/10.1016/j.pec.2010.01.013)] [Medline: [20202779](https://pubmed.ncbi.nlm.nih.gov/20202779/)]
5. Lin CT, Wittevrongel L, Moore L, Beaty BL, Ross SE. An Internet-based patient-provider communication system: randomized controlled trial. *J Med Internet Res* 2005;7(4):e47 [FREE Full text] [doi: [10.2196/jmir.7.4.e47](https://doi.org/10.2196/jmir.7.4.e47)] [Medline: [16236699](https://pubmed.ncbi.nlm.nih.gov/16236699/)]
6. Emont S. Measuring the impact of patient portals: What the literature tells us. 2011. URL: <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/M/PDF%20MeasuringImpactPatientPortals.pdf> [accessed 2013-06-11] [WebCite Cache ID 6HicVvNmP]
7. Leveille SG, Huang A, Tsai SB, Weingart SN, Iezzoni LI. Screening for chronic conditions using a patient internet portal: recruitment for an internet-based primary care intervention. *J Gen Intern Med* 2008 Apr;23(4):472-475 [FREE Full text] [doi: [10.1007/s11606-007-0443-6](https://doi.org/10.1007/s11606-007-0443-6)] [Medline: [18373147](https://pubmed.ncbi.nlm.nih.gov/18373147/)]
8. Horvath M, Levy J, L'Engle P, Carlson B, Ahmad A, Ferranti J. Impact of health portal enrollment with email reminders on adherence to clinic appointments: a pilot study. *J Med Internet Res* 2011;13(2):e41 [FREE Full text] [doi: [10.2196/jmir.1702](https://doi.org/10.2196/jmir.1702)] [Medline: [21616784](https://pubmed.ncbi.nlm.nih.gov/21616784/)]

9. Akesson KM, Saveman BI, Nilsson G. Health care consumers' experiences of information communication technology--a summary of literature. *Int J Med Inform* 2007 Sep;76(9):633-645. [doi: [10.1016/j.ijmedinf.2006.07.001](https://doi.org/10.1016/j.ijmedinf.2006.07.001)] [Medline: [16931133](https://pubmed.ncbi.nlm.nih.gov/16931133/)]
10. Neuhauser L, Kreps GL. eHealth communication and behavior change: promise and performance. *Social Semiotics* 2010 Feb;20(1):9-27. [doi: [10.1080/10350330903438386](https://doi.org/10.1080/10350330903438386)]
11. American Med Association. Will meaningful use spur the growth of patient portals?. URL: <http://www.ama-assn.org/amednews/2012/09/17/bisb0917.htm> [accessed 2012-11-16] [WebCite Cache ID [6CE5luOQx](https://www.webcitation.org/6CE5luOQx)]
12. Markle Foundation Initiative. Connecting for health - Americans overwhelmingly believe electronic personal health records could improve their health. URL: <http://www.policyarchive.org/handle/10207/bitstreams/15533.pdf> [accessed 2012-11-16] [WebCite Cache ID [6CE6UDNZV](https://www.webcitation.org/6CE6UDNZV)]
13. California HealthCare Foundation. Consumers and Health Information Technology: A National Survey. 2010. URL: <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/C/PDF%20ConsumersHealthInfoTechnologyNationalSurvey.pdf> [accessed 2012-11-16] [WebCite Cache ID [6CE6i8rFC](https://www.webcitation.org/6CE6i8rFC)]
14. Markle Foundation. Survey Finds Americans Want Electronic Personal Health Information to Improve Own Healthcare. URL: <http://www.policyarchive.org/handle/10207/bitstreams/15500.pdf> [accessed 2012-11-16] [WebCite Cache ID [6CE6zn5nD](https://www.webcitation.org/6CE6zn5nD)]
15. Deloitte Center for Health Solutions. Survey of Health Care Consumers: Key Findings, Strategic Implications. 2009. URL: http://www.deloitte.com/view/en_US/us/Insights/centers/center-for-health-solutions/5735e23a4b101210VgnVCM100000ba42f00aRCRD.htm [accessed 2012-11-16] [WebCite Cache ID [6CE7UiNDR](https://www.webcitation.org/6CE7UiNDR)]
16. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin CT. Expectations of patients and physicians regarding patient-accessible medical records. *J Med Internet Res* 2005;7(2):e13 [FREE Full text] [doi: [10.2196/jmir.7.2.e13](https://doi.org/10.2196/jmir.7.2.e13)] [Medline: [15914460](https://pubmed.ncbi.nlm.nih.gov/15914460/)]
17. Markle Foundation. Markle Survey on Health in a Networked Life, Snapshot: The Public and Doctors Agree with Blue Button Idea. URL: http://www.markle.org/sites/default/files/20101007_bluebutton_summary_1.pdf [accessed 2012-11-16] [WebCite Cache ID [6CE7fpuKS](https://www.webcitation.org/6CE7fpuKS)]
18. Walker J, Ahern DK, Le LX, Delbanco T. Insights for internists: "I want the computer to know who I am". *J Gen Intern Med* 2009 Jun;24(6):727-732 [FREE Full text] [doi: [10.1007/s11606-009-0973-1](https://doi.org/10.1007/s11606-009-0973-1)] [Medline: [19412641](https://pubmed.ncbi.nlm.nih.gov/19412641/)]
19. 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](https://doi.org/10.1111/j.1749-6632.2009.05337.x)] [Medline: [20201865](https://pubmed.ncbi.nlm.nih.gov/20201865/)]
20. Smedley BD. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: Board of Health Sciences Policy, Institute of Medicine; 2003.
21. Hussey P, Anderson G, Berthelot JM, Feek C, Kelley E, Osborn R, et al. Trends in socioeconomic disparities in health care quality in four countries. *Int J Qual Health Care* 2008 Feb;20(1):53-61 [FREE Full text] [doi: [10.1093/intqhc/mzm055](https://doi.org/10.1093/intqhc/mzm055)] [Medline: [18024997](https://pubmed.ncbi.nlm.nih.gov/18024997/)]
22. AHRQ. National Healthcare Disparities Report. 2011. URL: <http://www.ahrq.gov/research/findings/nhqrdr/nhdr11/index.html> [accessed 2013-03-17] [WebCite Cache ID [6FC6vIMNM](https://www.webcitation.org/6FC6vIMNM)]
23. Roundtable on Health Literacy - Institute of Medicine. Innovations in Health Literacy: Workshop Summary. Washington, DC: The National Academies Press; 2011.
24. Kutner MGE, Jin Y, Paulsen C. The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy. Washington DC: National Center for Education Statistics; 2006. URL: http://nces.ed.gov/pubs2006/2006483_1.pdf [accessed 2013-03-17] [WebCite Cache ID [6FC7KmFLH](https://www.webcitation.org/6FC7KmFLH)]
25. Rudd RE. Health literacy skills of U.S. adults. *Am J Health Behav* 2007;31 Suppl 1:S8-18. [doi: [10.5555/ajhb.2007.31.supp.S8](https://doi.org/10.5555/ajhb.2007.31.supp.S8)] [Medline: [17931141](https://pubmed.ncbi.nlm.nih.gov/17931141/)]
26. Nielsen-Bohman L, Panzer AM, Kindig DA. Health Literacy: A Prescription to End Confusion. Washington, DC: Institute of Medicine of the National Academies; 2004.
27. Schwartzberg JL, Van Geest JC, Wang CC. Understanding Health Literacy: Implications For Medicine And Public Health. Chicago: American Medical Association; 2005.
28. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med* 2004 Dec;19(12):1228-1239 [FREE Full text] [doi: [10.1111/j.1525-1497.2004.40153.x](https://doi.org/10.1111/j.1525-1497.2004.40153.x)] [Medline: [15610334](https://pubmed.ncbi.nlm.nih.gov/15610334/)]
29. Berkman NDSS, Donahue KE, Halpern DJ, Viera A, Crotty K, Holland A, et al. Health literacy interventions and outcomes: An updated systematic review. Evidence Report/Technology Assesment No. 199. (Prepared by RTI International--University of North Carolina Evidencebased Practice Center under contract No. 290-2007-10056-I. AHRQ Publication Number 11-E006. Rockville, MD: Agency for Healthcare Research and Quality; 2011. URL: <http://www.ahrq.gov/legacy/clinic/epcsums/litupsum.htm> [accessed 2013-03-17] [WebCite Cache ID [6FCAz5VpL](https://www.webcitation.org/6FCAz5VpL)]
30. Baker DW, Gazmararian JA, Williams MV, Scott T, Parker RM, Green D, et al. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health* 2002 Aug;92(8):1278-1283. [Medline: [12144984](https://pubmed.ncbi.nlm.nih.gov/12144984/)]
31. Davis TC, Williams MV, Marin E, Parker RM, Glass J. Health literacy and cancer communication. *CA Cancer J Clin* 2002;52(3):134-149. [Medline: [12018928](https://pubmed.ncbi.nlm.nih.gov/12018928/)]

32. Stableford S, Mettger W. Plain language: a strategic response to the health literacy challenge. *J Public Health Policy* 2007;28(1):71-93. [doi: [10.1057/palgrave.jphp.3200102](https://doi.org/10.1057/palgrave.jphp.3200102)] [Medline: [17363939](https://pubmed.ncbi.nlm.nih.gov/17363939/)]
33. Zarcadoolas C, Blanco M, Boyer JF, Pleasant A. Unweaving the Web: an exploratory study of low-literate adults' navigation skills on the World Wide Web. *J Health Commun* 2002;7(4):309-324. [doi: [10.1080/10810730290088157](https://doi.org/10.1080/10810730290088157)] [Medline: [12356289](https://pubmed.ncbi.nlm.nih.gov/12356289/)]
34. Kirsch I, Jungeblut A, Jenkins L, Kolstad A. *Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey (NALS)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education; 1993.
35. Chang BL, Bakken S, Brown SS, Houston TK, Kreps GL, Kukafka R, et al. Bridging the digital divide: reaching vulnerable populations. *J Am Med Inform Assoc* 2004;11(6):448-457 [FREE Full text] [doi: [10.1197/jamia.M1535](https://doi.org/10.1197/jamia.M1535)] [Medline: [15299002](https://pubmed.ncbi.nlm.nih.gov/15299002/)]
36. Dhanireddy S, Walker J, Reisch L, Oster N, Delbanco T, Elmore JG. The urban underserved: attitudes towards gaining full access to electronic medical records. *Health Expect* 2012 Jun 28:-. [doi: [10.1111/j.1369-7625.2012.00799.x](https://doi.org/10.1111/j.1369-7625.2012.00799.x)] [Medline: [22738155](https://pubmed.ncbi.nlm.nih.gov/22738155/)]
37. Schickedanz A, Huang D, Lopez A, Cheung E, Lyles CR, Bodenheimer T, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. *J Gen Intern Med* 2013 Jul;28(7):914-920. [doi: [10.1007/s11606-012-2329-5](https://doi.org/10.1007/s11606-012-2329-5)] [Medline: [23423453](https://pubmed.ncbi.nlm.nih.gov/23423453/)]
38. Patel VN, Dhopeswarkar RV, Edwards A, Barron Y, Likourezos A, Burd L, et al. Low-income, ethnically diverse consumers' perspective on health information exchange and personal health records. *Inform Health Soc Care* 2011 Dec;36(4):233-252. [doi: [10.3109/17538157.2011.554930](https://doi.org/10.3109/17538157.2011.554930)] [Medline: [21851182](https://pubmed.ncbi.nlm.nih.gov/21851182/)]
39. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. *J Gen Intern Med* 2011 Oct;26(10):1112-1116 [FREE Full text] [doi: [10.1007/s11606-011-1728-3](https://doi.org/10.1007/s11606-011-1728-3)] [Medline: [21538166](https://pubmed.ncbi.nlm.nih.gov/21538166/)]
40. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. *J Am Med Inform Assoc* 2009;16(5):683-689 [FREE Full text] [doi: [10.1197/jamia.M3169](https://doi.org/10.1197/jamia.M3169)] [Medline: [19567790](https://pubmed.ncbi.nlm.nih.gov/19567790/)]
41. 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/)]
42. 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/)]
43. Lyles CR, Harris LT, Jordan L, Grothaus L, Wehnes L, Reid RJ, et al. Patient race/ethnicity and shared medical record use among diabetes patients. *Med Care* 2012 May;50(5):434-440. [doi: [10.1097/MLR.0b013e318249d81b](https://doi.org/10.1097/MLR.0b013e318249d81b)] [Medline: [22354209](https://pubmed.ncbi.nlm.nih.gov/22354209/)]
44. 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/)]
45. Gustafson DH, Hawkins RP, Boberg EW, McTavish F, Owens B, Wise M, et al. CHES: 10 years of research and development in consumer health informatics for broad populations, including the underserved. *Int J Med Inform* 2002 Nov 12;65(3):169-177. [Medline: [12414016](https://pubmed.ncbi.nlm.nih.gov/12414016/)]
46. Czaja S, Charness N, Dijkstra K, Fisk A, Roger W, Sharit J. CREATE Common Core Battery measures. (CREATE-2006-01)
47. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychol Aging* 2006 Jun;21(2):333-352 [FREE Full text] [doi: [10.1037/0882-7974.21.2.333](https://doi.org/10.1037/0882-7974.21.2.333)] [Medline: [16768579](https://pubmed.ncbi.nlm.nih.gov/16768579/)]
48. US Department of Health and Human Services - Health Resources Services Administration. Find Shortage Areas: MUA/P by State and County. URL: <http://muafind.hrsa.gov/> [accessed 2013-03-17] [WebCite Cache ID [6FC9TPXP8](https://www.webcitation.org/6FC9TPXP8)]
49. Olson E, VanWye G, Kerker B, Thorpe L, Frieden T. Take Care East Harlem. NYC Community Health Profiles 2006;21(42):1-16 [FREE Full text]
50. Olson E, VanWye G, Kerker B, Thorpe L, Frieden T. Take Care Northwest Queens. NYC Community Health Profiles 2006;29(42):1-16 [FREE Full text]
51. US Census Bureau. 2010 Census, SF1, Population Division - New York City Department of City Planning. URL: <http://maps.nyc.gov/census/> [accessed 2013-07-29] [WebCite Cache ID [6IUKUeOBR](https://www.webcitation.org/6IUKUeOBR)]
52. Olson E, VanWye G, Kerker B, Thorpe L, Frieden T. Take Care Flatbush. NYC Community Health Profiles 2006;14(42):1-16 [FREE Full text]
53. Office of Institutional Advancement and the Office of Planning SDMC. Report on Brooklyn Community Health. 2007. URL: <http://www.downstate.edu/bhr/reports/Community-Health-Report.pdf> [accessed 2013-03-17] [WebCite Cache ID [6FCAmbuTx](https://www.webcitation.org/6FCAmbuTx)]
54. Abrahams JP, Hoyer WJ, Elias MF, Bradigan B. Gerontological research in psychology published in the *Journal of Gerontology* 1963-1974: Perspectives and progress. *J Gerontol* 1975 Nov;30(6):668-673. [Medline: [1184925](https://pubmed.ncbi.nlm.nih.gov/1184925/)]

55. Killoran AHK, Dalley G. Promoting the health of older people: a compendium.: Health Education Authority; 1997. URL: http://www.gserve.nice.org.uk/nicemedia/documents/promohealth_older_comp.pdf [accessed 2013-03-17] [WebCite Cache ID 6FCArUKTO]
56. Lake Research Partners. National Survey of Adults 50 and Older on Experiences with Health Care System. URL: http://www.nationalpartnership.org/site/DocServer/Lake_Poll_Media_Report_FINAL.pdf?docID=6242 [accessed 2013-03-17] [WebCite Cache ID 6FC6XEVQo]
57. Glaser B. The constant comparative method of qualitative analysis. *Social Problems* 1965;12:436-445.
58. Glaser B. *Basics of Grounded Theory Analysis: Emergence Vs. Forcing*. Mill Valley, CA: Sociology Press; 1992.
59. Holton J. The coding process and its challenges. In: Bryant A, Charmaz K, editors. *Sage Handbook of Grounded Theory*. London: Sage Publications; 2007:265-289.
60. Kim S, Egerter S, Cubbin C, Takahashi ER, Braveman P. Potential implications of missing income data in population-based surveys: an example from a postpartum survey in California. *Public Health Rep* 2007;122(6):753-763 [FREE Full text] [Medline: 18051668]
61. Winkleby MA, Cubbin C. Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States. *J Epidemiol Community Health* 2003 Jun;57(6):444-452 [FREE Full text] [Medline: 12775792]
62. Fox S. Health topics. American: Pew Internet & Life Project; 2011. URL: <http://www.pewinternet.org/Reports/2011/HealthTopics.aspx> [accessed 2012-11-16] [WebCite Cache ID 6CE7rqUog]
63. Vance K, Howe W, Dellavalle RP. Social internet sites as a source of public health information. *Dermatol Clin* 2009 Apr;27(2):133-6, vi. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]
64. Chou WY, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. *J Med Internet Res* 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
65. Fox S, Duggan M. Mobile Health: Half of smartphone owners use their devices to get health information and one-fifth of smartphone owners have health apps.: Pew Internet Research; 2012. URL: http://pewinternet.org/~media/Files/Reports/2012/PIP_MobileHealth2012.pdf [accessed 2013-05-31] [WebCite Cache ID 6H1DjC4zR]
66. Purcell K, Entner R, Henderson N. The Rise of Apps Culture. 2010. URL: <http://pewinternet.org/Reports/2010/The-Rise-of-Apps-Culture.aspx> [accessed 2013-05-31] [WebCite Cache ID 6H1EXqsIA]
67. Luque AE, van Keken A, Winters P, Keefer MC, Sanders M, Fiscella K. Barriers and Facilitators of Online Patient Portals to Personal Health Records Among Persons Living With HIV: Formative Research. *JMIR Res Protoc* 2013;2(1):e8. [doi: 10.2196/resprot.2302] [Medline: 23612564]
68. Hilton JF, Barkoff L, Chang O, Halperin L, Ratanawongsa N, Sarkar U, et al. A cross-sectional study of barriers to personal health record use among patients attending a safety-net clinic. *PLoS One* 2012;7(2):e31888 [FREE Full text] [doi: 10.1371/journal.pone.0031888] [Medline: 22363761]
69. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. *J Med Internet Res* 2013;15(2):e43 [FREE Full text] [doi: 10.2196/jmir.2243] [Medline: 23470453]
70. Joyce CR, Caple G, Mason M, Reynolds E, Mathews JA. Quantitative study of doctor-patient communication. *Q J Med* 1969 Apr;38(150):183-194. [Medline: 5781349]
71. Roter DL, Hall JA. Studies of doctor-patient interaction. *Annu Rev Public Health* 1989;10:163-180. [doi: 10.1146/annurev.pu.10.050189.001115] [Medline: 2655629]
72. Anderson JL, Dodman S, Kopelman M, Fleming A. Patient information recall in a rheumatology clinic. *Rheumatol Rehabil* 1979 Feb;18(1):18-22. [Medline: 311507]
73. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealthVet Personal Health Record Pilot: qualitative study. *J Med Internet Res* 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
74. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, et al. Improving diabetes management with a patient portal: a qualitative study of diabetes self-management portal. *J Med Internet Res* 2012;14(6):e158 [FREE Full text] [doi: 10.2196/jmir.2265] [Medline: 23195925]

Abbreviations

- CAMBA:** a community-based organization in Brooklyn committed to serving low-income, homeless, and immigrant populations, among other vulnerable groups.
- GT:** grounded theory
- HITECH:** Health Information Technology for Economic and Clinical Health Act
- HRSA:** Health Resources and Services Administration
- IRB:** Institutional Review Board

Edited by G Eysenbach; submitted 20.12.12; peer-reviewed by YM Tan, A Hörbst; comments to author 20.01.13; revised version received 18.03.13; accepted 13.06.13; published 26.08.13.

Please cite as:

Zarcadoolas C, Vaughn WL, Czaja SJ, Levy J, Rockoff ML
Consumers' Perceptions of Patient-Accessible Electronic Medical Records
J Med Internet Res 2013;15(8):e168
URL: <http://www.jmir.org/2013/8/e168/>
doi: [10.2196/jmir.2507](https://doi.org/10.2196/jmir.2507)
PMID: [23978618](https://pubmed.ncbi.nlm.nih.gov/23978618/)

©Christina Zarcadoolas, Wendy L Vaughn, Sara J Czaja, Joslyn Levy, Maxine L Rockoff. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 26.08.2013. 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

Can Social Semantic Web Techniques Foster Collaborative Curriculum Mapping In Medicine?

Cord Spreckelsen¹, Dr rer nat; Sonja Finsterer², Dr med; Jan Cremer^{2,3}, Dr med; Hennig Schenkat^{1,2}

¹Division of Knowledge-Based Systems, Department of Medical Informatics, RWTH Aachen University, Aachen, Germany

²Deanery of Student Affairs, Faculty of Medicine, RWTH Aachen University, Aachen, Germany

³Department of Anesthesiology and Intensive Care Medicine, RWTH Aachen University, Aachen, Germany

Corresponding Author:

Cord Spreckelsen, Dr rer nat

Division of Knowledge-Based Systems

Department of Medical Informatics

RWTH Aachen University

Pauwelsstr 30

Aachen, 52057

Germany

Phone: 49 241338088870

Fax: 49 241338088870

Email: CSpreckelsen@mi.rwth-aachen.de

Abstract

Background: Curriculum mapping, which is aimed at the systematic realignment of the planned, taught, and learned curriculum, is considered a challenging and ongoing effort in medical education. Second-generation curriculum managing systems foster knowledge management processes including curriculum mapping in order to give comprehensive support to learners, teachers, and administrators. The large quantity of custom-built software in this field indicates a shortcoming of available IT tools and standards.

Objective: The project reported here aims at the systematic adoption of techniques and standards of the Social Semantic Web to implement collaborative curriculum mapping for a complete medical model curriculum.

Methods: A semantic MediaWiki (SMW)-based Web application has been introduced as a platform for the elicitation and revision process of the Aachen Catalogue of Learning Objectives (ACLO). The semantic wiki uses a domain model of the curricular context and offers structured (form-based) data entry, multiple views, structured querying, semantic indexing, and commenting for learning objectives ("LOs"). Semantic indexing of learning objectives relies on both a controlled vocabulary of international medical classifications (ICD, MeSH) and a folksonomy maintained by the users. An additional module supporting the global checking of consistency complements the semantic wiki. Statements of the Object Constraint Language define the consistency criteria. We evaluated the application by a scenario-based formative usability study, where the participants solved tasks in the (fictional) context of 7 typical situations and answered a questionnaire containing Likert-scaled items and free-text questions.

Results: At present, ACLO contains roughly 5350 operational (ie, specific and measurable) objectives acquired during the last 25 months. The wiki-based user interface uses 13 online forms for data entry and 4 online forms for flexible searches of LOs, and all the forms are accessible by standard Web browsers. The formative usability study yielded positive results (median rating of 2 ("good") in all 7 general usability items) and produced valuable qualitative feedback, especially concerning navigation and comprehensibility. Although not asked to, the participants (n=5) detected critical aspects of the curriculum (similar learning objectives addressed repeatedly and missing objectives), thus proving the system's ability to support curriculum revision.

Conclusions: The SMW-based approach enabled an agile implementation of computer-supported knowledge management. The approach, based on standard Social Semantic Web formats and technology, represents a feasible and effectively applicable compromise between answering to the individual requirements of curriculum management at a particular medical school and using proprietary systems.

(*J Med Internet Res* 2013;15(8):e169) doi:[10.2196/jmir.2623](https://doi.org/10.2196/jmir.2623)

KEYWORDS

curriculum mapping; medical education; Semantic Web; Social Web

Introduction

Background

Curriculum mapping supports teachers, learners, and curriculum administrators by providing a comprehensive overview of a curriculum and its elements and their interrelations. It answers questions like “Where do we teach what?” [1]. Curriculum mapping is considered to be a demanding, data-intensive, and essentially collaborative effort [2]. The number of custom-built software platforms for curriculum mapping indicates a lack of available tools and standards [3]. Social Semantic Web (SSW) approaches combine Web content, which can be partly “understood” (ie, processed in a semantically sound way) by computer programs with social software, enabling the collaborative creation and maintenance of content to take place. Thus, SSW approaches provide promising solutions for implementing and maintaining curriculum mapping in medicine as a knowledge management process.

The revision of the national German medical licensing regulations enacted in 2002 enabled markedly changed medical curricula to be introduced at German medical schools, and this included the creation of a number of reform model curricula [4]. This change process increased the need for computer-supported curriculum mapping.

Curriculum Mapping

In the 1970s, Hausman coined the term curriculum mapping in the context of curriculum planning [5]. Describing the differences between the prescribed, the taught, and the tested curricula (ie, the ideal curriculum as planned in advance, the curriculum delivered by the teachers, and the curriculum learned by the students), English then proposed curriculum mapping as an approach to realign these three “circles”, mainly by capturing the taught curriculum and comparing it to the ideal one [6]. Gjerde stressed the potentially positive effect of curriculum mapping on the congruence of learning objectives and tests used for evaluation [7].

In his comprehensive review of curriculum mapping, Harden recommended the approach as a pivotal factor in fostering coordination and communication of medical curricula [8]. He also emphasized the aspect of depicting the components of a curriculum and their interrelations explicitly in a curriculum map, which should provide different views (“windows”) focusing on, for example, the content, the learning events, or the learning resources of a given curriculum.

In 2008, a survey on the status of curriculum mapping in Canada and the United Kingdom found that 55% of the responding medical schools were in the process of establishing a curriculum map [3]. Notably, Willett also found that curriculum mapping was considered to be an ongoing process requiring “continual upgrading and maintenance”. Following their analysis, the large quantity of custom-built software indicates a shortcoming of available IT tools that meet the requirements of curriculum management.

Computer-Supported Curriculum Management

The Faculty of Medicine at McGill University implemented an electronic curriculum map at an early stage; this allows a curriculum inventory to be performed [9]. Tufts Health Science Database (HSDB) was an early attempt to integrate content delivery and curriculum management. Lee et al describe the positive effects of HSDB on faculty development, curricular reform, and interdisciplinarity [10]. CurrMIT serves as a means to capture, manage, and compare the curricula of North American medical schools [11]. Curriculum mapping tools have also been developed in other health-related areas, for example, in the context of nursing education [12].

In contrast to the awareness of the role of electronic curriculum maps for innovative curriculum management and the early attempts to integrate learning content, many approaches—especially in German-speaking countries—focused on creating online learning objectives catalogs or databases. The Swiss Catalogue of Learning Objectives (SCLO) [13] was one of the most influential European projects for defining and managing structured, outcome-oriented learning objectives. SCLO provides an open-access Web portal that allows the objectives to be filtered by keyword, type, discipline, topic, and competence levels. In the meantime, LO catalogs similar to the SCLO have been implemented at various German medical schools [14–16]. The projects grant Web-based access to their users. All Web platforms are custom-built, which corresponds to the general trend found by Willett at UK and Canadian medical schools [3]. The online tool used for the Heidelberg Catalogue of Learning Objectives (HCLO) focuses on enabling interactive maintenance and improvement of the catalog. The Charité University Hospital, Berlin, uses its platform to support systematic curriculum mapping, and claims to be one of the first German faculties to do so. Hege et al similarly advocated the use of computer-based LO catalogs in the process of managing and mapping a curriculum [15,16]. In addition to catalogs of learning objectives, which address the complete curricula of medical schools, some approaches concentrate on a given medical specialty or field [17].

Second-Generation Curriculum Management Systems

Watson et al distinguished a first from a second generation of curriculum management systems: the first generation comprises electronic curriculum maps and databases that support administrative processes, whereas the second generation is represented by “comprehensive knowledge-management systems primarily designed to support students and teachers in the learning and teaching process while also supporting administrative processes” [18].

eMed, developed at the University of New South Wales, is reported to be an integrated second-generation curriculum management system, combining various tools and services. eMed supports very different aspects of curriculum management, ranging from curriculum mapping to student portfolios, and bridging the gap between organizational, curricular, and economic needs [18]. Bell et al reported on a

curriculum-mapping project that concentrates especially on curricular quality management by systematic curriculum review; this fosters integration, transparency, and communication [19]. The authors found the curriculum mapping process to be resource-intensive, and underlined the fact that none of the existing software tools designed for supporting curriculum mapping fully met the requirements of their project. Just recently, the University of Toronto developed CMap, a special computer-based curriculum-mapping tool that was able to detect an uneven distribution of teaching time with respect to the major topics and skills of a given planned curriculum [20].

Semantic Indexing

An important prerequisite for enabling curriculum mapping and comprehensive curriculum management, based on databases organizing curricular data and learning objective catalogs, is semantic indexing. Learning objectives need to be retrievable by medical topic, similar learning content should be associated, and the coverage of learning objectives by learning events or examinations needs to become visible.

Denny et al investigated the ability of a text mining and classification tool to give an automatic estimate of the coverage of medical topics by lectures, based on texts documenting these lectures [21]. The approach proved successful in supporting semantic indexing of curricular events and outperformed, in that curricular context, a similar tool (MetaMap) that had primarily been designed for indexing scientific publications [21,22]. The taxonomy TIME (Topics for Indexing Medical Education) served as a means to index the elements of a curriculum map uniformly [23]. Thus, TIME enabled topic specific views, which show the contribution of curricular elements to specific outcomes. Dexter et al reported on using the US Medical Licensing Examination Step 1 Content Outline (USMLE Step1 CO) as a basis for indexing LOs in order to assess the completeness of topic coverage [24]. Instead of using preexisting taxonomies, semantic indexing can also be achieved by collaboratively created taxonomies, which evolve during the use of a given information system, so-called folksonomies. Gasevic et al reported an exemplary application of folksonomies to the maintenance of learning environments [25].

Aachen Catalogue of Learning Objectives

In 2003, the Medical Faculty of the RWTH Aachen University started the Aachen Medical Model Curriculum (AMMC). The curriculum implements a spiral-shaped education process ("spiral curriculum" [26]) and aims at an early integration of preclinical and clinical education. In the second and third years, 11 multidisciplinary modules, each focusing on a particular organ system, form the backbone of the curriculum. The AMMC follows an explicit mission statement and defines learning goals for all modules. Furthermore, a faculty-spanning consensus process, which took place before the start of the model curriculum, defined the content of all modules. Nonetheless, the documentation of learning objectives merely relied on item lists, which proved insufficient for maintaining and further developing the curriculum. Thus, in 2011, the faculty formally decided to implement a comprehensive, Web-based catalog of competency-based learning objectives (Aachen Catalogue of Learning Objectives or ACLO). The spiral curriculum and

multidisciplinary modules create challenging requirements concerning the elicitation, revision, and communication process of the LOs.

Aims

The Web-based Aachen Catalogue of Learning Objectives (ACLO-Web) aims to support the implementation, maintenance, and use of the ACLO by establishing a Web-based knowledge management system based on Social Semantic Web technology and standards.

Methods

Elicitation and Revision Process of the ACLO

The faculty formed a special task force ("Learning Objectives Working Group") in order to organize the implementation of the ACLO. The dean asked all contributing units (clinics, institutes, and external departments) to name representatives responsible for the detailed specification of LOs. Each representative needed to participate in a training program: mainly a 1-day workshop by trainers with certified competency in medical education. After the training, the representatives started the elicitation and specification of LOs based on predefined forms, while consulting everyone in their unit (clinic or institute) who was involved in teaching. In order to guarantee a high and equal level of data quality, especially with respect to thematic indexing of the LOs, a team of 1 medical expert and 2 student coworkers (the "ACLO Team") was responsible for the primary data entry into the central catalog. This team continuously checked and improved the assignment of LOs to medical topics, curricular modules, and a responsible faculty. Furthermore, they supervised quality standards concerning the formulation of competency-based learning objectives. During the ACLO revision, the responsible representative, who could of course involve his or her colleagues, checked and eventually improved the LO specification.

When the Learning Objectives Working Group finally approved the LOs, the catalog was made accessible to students. Faculty members and students now continuously check the LO catalog for inconsistencies and monitor whether the LOs are adequately addressed by the courses and lectures. Their feedback should lead to the future improvement of the catalog.

Requirements

An information system supporting the process described below has to meet the following requirements:

- Collaborative: The system needs to support a collaborative effort. It needs to provide decentralized access to a central LO repository, allow the storage and tracking of the whole history of changes made to the LOs in the collection (ie, support a versioning mechanism), and manage user accounts and roles.
- Structured: The system must allow the structured entry and retrieval of information; in order to enable a systematic search of different aspects of the LO specification, the system needs to store structured information (attribute-value pairs of the different aspects of the LO specification), which can be entered in online forms.

- **Flexible:** The system needs to be flexible with respect to changes in the representation, retrieval, and presentation of LOs. Ideally, the system should support the implementation process of the LO catalog from the very beginning. As a consequence of the ongoing implementation process, it should be possible to adapt the system to changing requirements.

Social Semantic Web-Based Curriculum Management

The online catalog of learning objectives is based on Social Semantic Web technology, in order to support collaborative knowledge management flexibly. The project reported here uses MediaWiki (the software platform of Wikipedia) enhanced by additional software modules (extensions). We adopted the Semantic MediaWiki (SMW) and Semantic Forms extensions.

A MediaWiki application enables the collaborative maintenance of Web pages; the pages can easily be edited and linked by wiki users using only a Web browser. Wiki users can revisit or even undo all changes to a page logged by the system in the page history. Similar content can be presented uniformly by using wiki templates combining predefined text fragments with new information entered via variables. Comments can be added to existing wiki pages. MediaWiki provides different approaches for managing user accounts and granting access. At present, read access to ACLO requires a general password communicated to all students and teachers. Readers can also enter comments and annotations (reader role). ACLO grants write access (ie, the right to change the core content), to registered users only (author role). The creation of user accounts is restricted to the administrators of ACLO.

As a first step toward equipping the wiki pages with computer-readable semantics, pages can be assigned to categories, and pages addressing similar concepts can be organized together (eg, the categories “disease” and “symptom” may be defined for pages addressing diseases and symptoms, respectively). Based on such categories, MediaWiki automatically generates index pages.

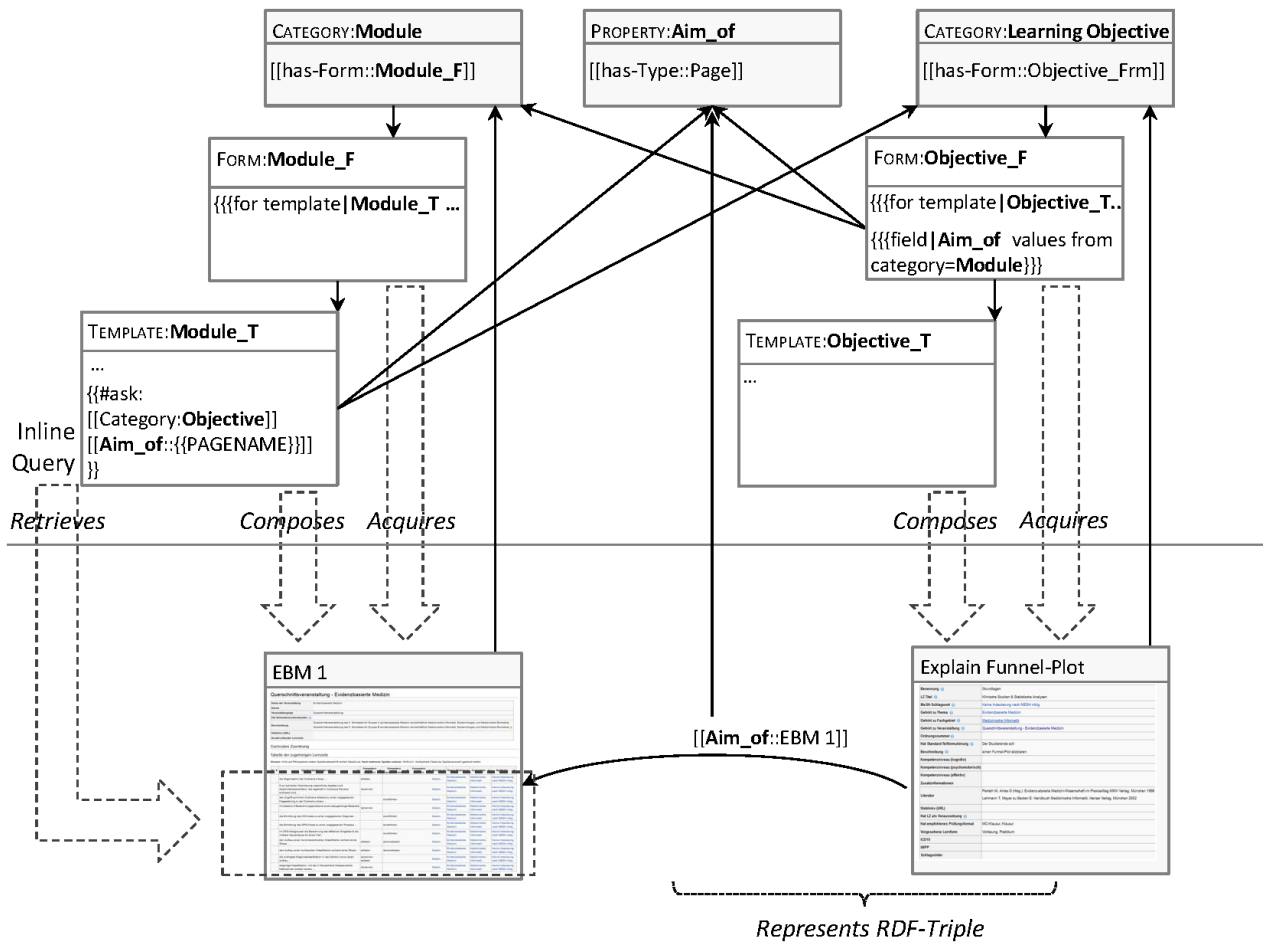
The SMW extension [27] adds semantic relations (SMW properties) to the MediaWiki, allowing pages to be annotated by structured property-value pairs and associated by meaningful relations. A page describing a learning objective may refer (link) to a different page describing a teaching module by the property “Aim_of” (syntactically achieved by inserting the SMW annotation “[[Aim_of::*target-page*]]” on the referring wiki page) (see Figure 1). SMW properties form subject-predicate-object triples (referring page, property, referenced page, or value) as introduced by the Semantic Web approach [28]. The enhanced representational means of SMW directly correspond to a Web Ontology Language (OWL) based knowledge representation: each SMW page represents an (abstract) individual, SMW properties correspond to OWL properties, and wiki categories

to OWL classes, respectively. In general, SMW annotations can be translated—and actually exported—to OWL DL statements (in OWL/RDF encoding), where OWL DL is an OWL sublanguage formally based on complete and decidable description logics [29,30]. Only some built-in SMW properties lack direct equivalents in OWL and are thus treated as annotation properties [27]. According to the above correspondence, most SMW annotations represent assertional knowledge, that is, they state facts concerning attributes of domain objects and their mutual relationships (A-Box statements of an ontology-based knowledge representation). Furthermore, but to a lesser extent (see below), SMW is able to represent statements defining domain concepts, for example, concerning classes of objects and admissible values of object attributes belonging to a given class, thus forming a conceptual schema of a domain (T-Box statements). Some SMW applications exploited this further by importing existing ontologies.

SMW also defines an elaborated query syntax operating on its semantic annotations. SMW pages may include inline queries enabling the dynamic and consistent updating of facts throughout the wiki application. Furthermore, it is possible to not only query, but also to manage an SMW application based on a Resource Description Framework database (RDF-Triple Store) in combination with the SPARQL Protocol and RDF Query Language (SPARQL) [27,31].

SMW properties may also refer to primitive data types, such as strings, numbers, or dates, allowing pages to organize structured data instead of containing unstructured text. The Semantic Forms extension then supports structured data entry by defining forms. Thus, forms can foster structured and uniform data entry for all pages of a given category. When authors enter content via semantic forms, the created wiki pages are categorized automatically. Semantic forms offer drop-down lists, check boxes, and a sophisticated auto-completion mechanism, and allow data entry to be restricted to semantically sound values. As shown in Figure 1, the approach relies mainly on semantic forms when imposing schema constraints. While it is possible to define all categories and properties necessary for representing a given field, the declaration of a property can use only very general data types and does not allow its values to be restricted to a domain defined in terms of existing categories (as a T-Box of a domain ontology would do). The semantic forms are the means available for imposing the respective constraints. Figure 1 shows the SMW mechanisms imposing semantic constraints and enabling structured data management; each SMW category can use a semantic template for uniformly composing a page out of structured parameters and each template can make use of semantic forms allowing controlled data entry by online forms. Most notably, a template can include query syntax that dynamically and individually refers to the content of each page composed by that template.

Figure 1. The SMW-based approach to the representation and acquisition of curricular information, illustrated by an example taken from ACLO-Web; the diagram includes abbreviated SMW syntax in order to illustrate why the references between the different building blocks are established.



caption: —→ Reference - - - - -> SMW functionality

Core Data Model

Figure 2 shows the core data model of the LO catalog. The diagram focuses on the most relevant aspects: each LO needs to be assigned to teaching modules of the curriculum (ie, courses, lectures, practicals, etc). Additionally, an LO needs to be characterized by the medical topic addressed by the LO and the medical specialty involved. The left-hand area of the diagram represents the curricular context (given by “Term”, “Segment”, “Curriculum”). The regulations of the course of study and the examinations define a planned curriculum. As official regulations are revised from time to time, teaching modules may be redefined or moved to different semesters. Thus, the catalog can reuse LOs in different modules following different regulations by assigning them to modules belonging to different versions of the curriculum.

Assigning a medical specialty to a teaching module implies that the respective clinic or institute is responsible for teaching (is “involved_in”) the multidisciplinary module. The direct relation “defines” indicates (institutional) authorship of medical objectives, which are learning objectives specifically addressing medical topics. The vast majority of LOs of ACLO belong to this type and are authored by teachers affiliated with a clinic or institute. The remaining LOs (some more general learning goals)

were defined by conferences. These LOs are not associated with specific institutions.

The model represents medical topics (eg, diseases, therapeutic approaches, or professional skills) as independent entities. Therefore, LOs can be assigned to a set of medical topics, which nonetheless may be revised or extended during the implementation process of the catalog. At present, all authors of learning objectives can enter new topics, while an auto-completion function reduces the risk of entering syntactic variants and fosters the user’s awareness of existing topics. Users can also form hierarchic or associative links between topics. Thus, the set of available topics and its structure evolves due to collaborative effort and, therefore, forms a folksonomy. The assignment of topic to LOs is carried out manually. It complements the automatic categorization of wiki content by the semantic forms. The use of a controlled set of topics promotes consistent semantic indexing and retrieval of LOs. The model also indicates references to external information. By referencing a MeSH-ID, an ICD-Code, an OPS-Code, or an IMPP-ID, the topics are linked to established medical classifications. These references establish links to the Medical Subject Headings (MeSH), the International Classification of Diseases (ICD), the German modification of the International Classification of Procedure in Medicine (Operationen- und

Prozedurenschlüssel - OPS), and the catalog of the Institute for Medical and Pharmaceutical Examination Items (IMPP), the national provider for final medical examinations, respectively. The system therefore allows it to be determined which requirements of the final examinations are covered by the learning objectives of a given module, term, and the whole curriculum, respectively. Additionally, the indexing of learning objectives by topics allows similar learning objectives of a given topic to be retrieved, which is achieved at present by using SMW subqueries within the system.

Additionally, the model provides 2 interfaces to different applications established at the medical school: (1) the SAM-ID leads to a semantic network used to index learning media produced by the local center for audiovisual media, and (2) the mediCal-ID links to the online calendar and room management system used for planning the terms. Both systems will be further integrated in the near future.

Specification of Learning Objectives

The assignment of LOs to teaching modules, topics, and responsible specialties (treated as independent model entities) is complemented by a detailed specification of each LO. Roughly, this specification includes (1) the essential description of the LO, (2) further indexing, and (3) supplementary didactic information.

The essential LO description formulates the observable behavior of the students, showing the level of competence required by the LO. Each LO description begins with a text introduction, which is chosen from a given set of templates (eg, “The students are able to...”, “At least the top 10% of the participants can...”). The sentence is continued by free text describing the required behavior and then finished by a verb (to be selected out of a given set) indicating the level of competence (eg, “explain”, “show”, “appraise”), in a sequence matching German grammar. Semantic indexing is enabled by assigning predefined topics as described above. Supplementary didactic information allows

Textbox 1. Criterion C1.

```
context curriculum inv: self.belong_to implies count(select(f|f.belongs_to.defined_in= self.belongs_to.defined_in and f.follows> size()-0))=1
```

Formative Usability Study

A formative usability study assists the implementation of an SMW-based online catalog. Formative usability studies can be successfully carried out with only a few participants who can nonetheless indicate the most relevant usability problems. Nielsen recommended 5-8 testers [28].

We chose a scenario-based approach. The participants used the online catalog, sequentially solving tasks given in the context of a typical scenario that they were asked to imagine. We defined 7 imaginary scenarios involving both students and medical teachers:

- (S1) A student preparing for an examination and retrieving the learning objectives of courses recently visited
- (S2) A student interested in a topic because a relative suffers from a particular disease

an LO to be associated with other LOs that are named as prerequisites for achieving the LO in question (predecessor LOs). It is also possible to indicate recommended learning and teaching formats for the LO (eg, “Lecture”, “Problem-based learning”) and recommended assessment formats (eg, “Multiple choice test”, “Objective structured practical examination”).

Global Consistency Checking

SMW technology enables structured data entry, semantic queries, and the dynamic consistent update of a large hypertext. Nonetheless, due to the relatively weak schema constraints imposed by the definition of categories and the semantic properties, complex or global semantic consistency criteria cannot be enforced algorithmically. As an example, the SMW platform alone provides no means for checking if an LO that is declared to be a prerequisite of a given LO is associated with prior learning events. In addition, complex cardinality constraints cannot be imposed on the SMW properties (eg, “There should be exactly one study section in each curriculum that has no predecessor”; in what follows, this is referred to as C1).

Therefore, ACLO-Web complements the SMW platform by adding a component enabling the definition and algorithmic check of enhanced consistency criteria (the ACLO consistency module or ACLO-CM). ACLO-CM is designed as a separate Web application operating on the SMW Triple Store and an additional constraint repository (Figure 3).

Consistency criteria are represented by expressions of the Object Constraint Language (OCL). Criterion C1, introduced above, is shown in Textbox 1.

ACLO-CM enables global consistency checking to be carried out at defined milestones. The component is not designed to prompt the users directly during data entry (as this would not be feasible for global constraints). Instead, it produces a detailed checklist, to be used for revision, of the problems found.

- (S3) A student trying to gain an overview of the following term
- (S4) A student planning to specialize in a given medical field in the future, who is interested in where his or her chosen field is present in the curriculum
- (S5) A medical teacher preparing for a multidisciplinary course
- (S6) A medical teacher planning an examination
- (S7) A medical teacher trying to find out about the prior knowledge of the students enrolled in his or her course.

Data was acquired using an online questionnaire containing the textual description (vignette) of the scenarios and instructing the participants on the specific tasks for each scenario. After each scenario, the participants were asked to rate the following statements using a Likert scale: (1) “I could intuitively carry out all necessary actions”, (2) “I found all relevant information”, and (3) “There was too much redundant information”. Items 2

and 3 were accompanied by a free-text field where redundant and missing information could be further specified.

General feedback was acquired by 8 Likert-scaled items concerning (1) the relevance of the LOs retrieved, (2) the completeness of the results, (3) the performance of the system, (4) the feasibility of a meaningful interpretation of the results, (5) the comprehensibility of the labels used by the system, (6) the structure of the system's output, (7) the design of the graphical user interface, and (8) the system's usability in general. Finally, the participants could enter free-text feedback on positive aspects, negative aspects, and recommendations for improvements.

Methods of Analysis

Due to the small number of participants, the scaled items of the usability test were analyzed by descriptive statistics only. The median, first quartile, minimal, and maximal values were derived

from the data and visualized by box plot diagrams. The free-text feedback was analyzed by bottom-up qualitative text coding. Suitable keywords (codes) were produced and assigned to the text of the users' statements while reading the text for the first time, the set of keywords was ordered and normalized, and then the keywords were used to index the statements consistently while reading the text for the second time. The statements were then rearranged by producing a synopsis of statements assigned to the same keywords and interpreted.

Software

The online questionnaire for the usability study was based on LimeSurvey (v 1.85) [32]. Descriptive statistics and box plots were produced by R (v 2.14.0) [33]. Qualitative text analysis used Microsoft OneNote 2010. The online catalog was implemented on a virtualized Debian Linux-based XAMP server (v 1.8.0) [34] using MediaWiki (v 1.18.5 with SemanticBundle extension r20120327) [35].

Figure 2. Core data model of ACLO-Web (grey boxes correspond to SMW categories, arrows correspond to SMW semantic relations; arrow subscripts (1-1, 1-n, n-m) correspond to one-to-one, one-to-many, and many-to-many relationship types, respectively; and boxes with rounded corners indicate references to external classification systems established by ID fields).

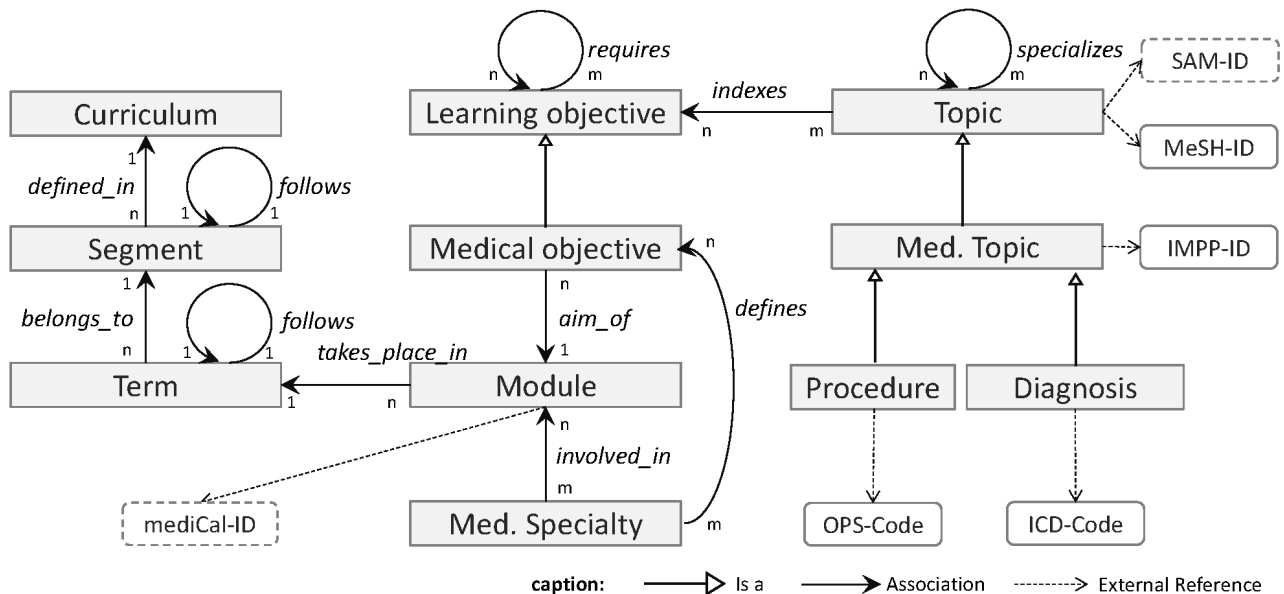
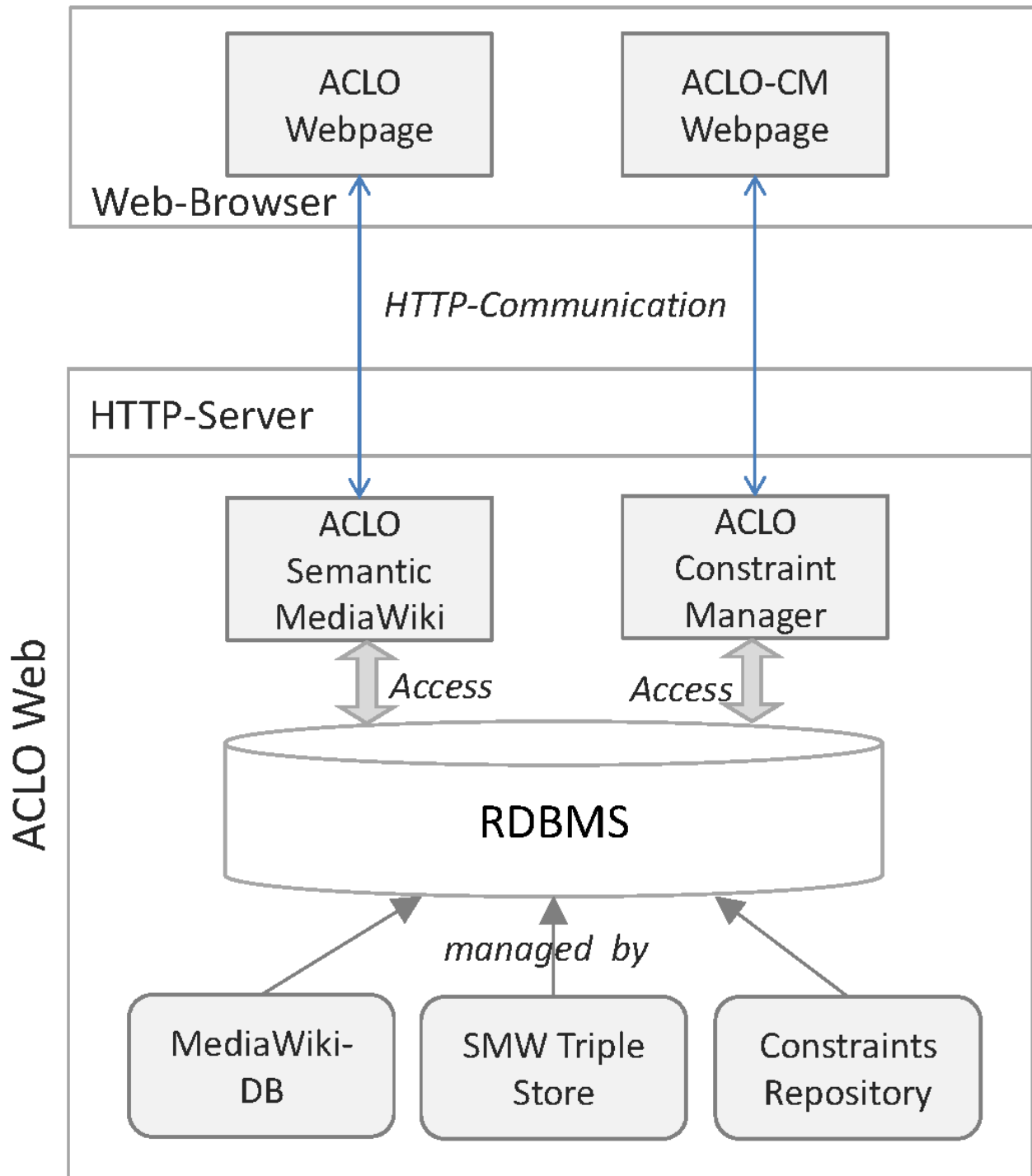


Figure 3. Architecture of ACLO-Web.

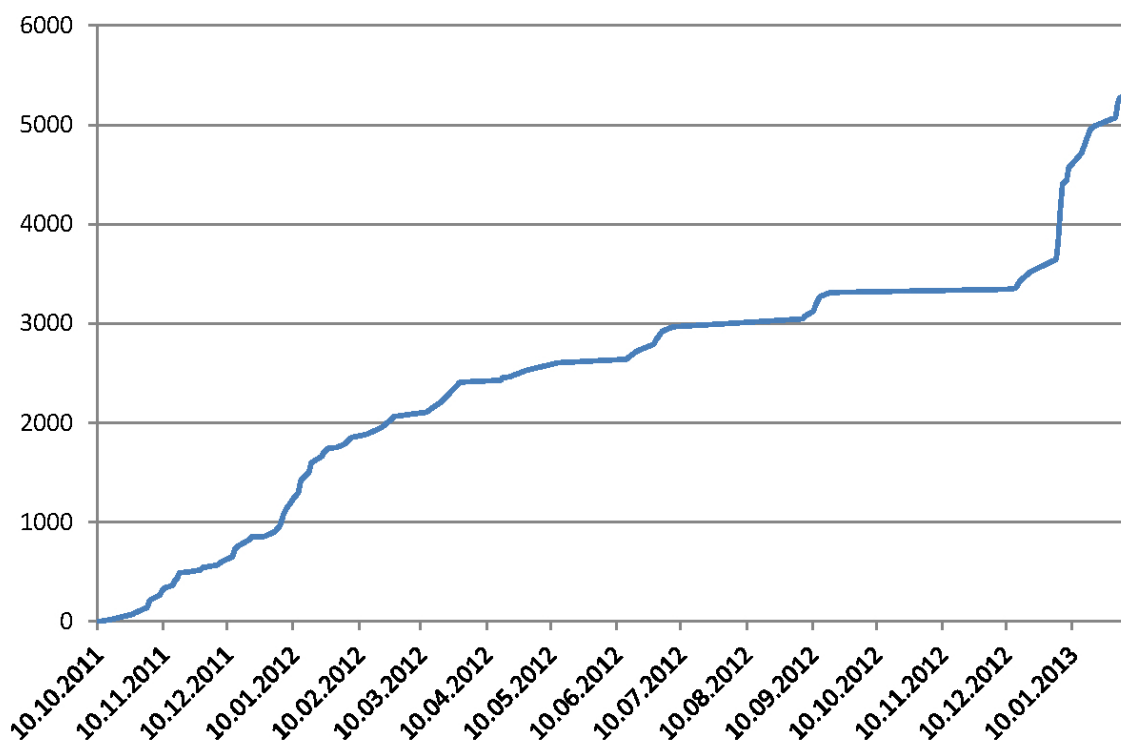


Results

State of ACLO-Web

The online catalog of learning objectives is based mainly on 9 wiki categories (“Code/ID”, “Curriculum”, “Teaching Module”, “Learning Objective”, “Medical Learning Objective”, “Medical Specialty”, “Medical Topic”, “Segment”, and “Term”), 42 wiki

templates, and 49 semantic properties of the SMW. The main activity concerning the data entry of LOs started in September 2011. By February 2013, 5350 LOs of the Aachen Medical Model Curriculum had been collected. The LOs were assigned to 69 modules, 61 medical specialties, and 243 medical topics that are defined in the system. Figure 4 shows the growth of the number of acquired LOs during this time interval.

Figure 4. State of ACLO-Web: date vs total number of Learning Objectives acquired.

Lessons Learned During Implementation

The use of the free, open source SMW platform enabled a fast and low-budget implementation of ACLO-Web. A first version was available—and immediately used by the LO team—only a few weeks after starting the project. Therefore, the system (categories, attributes, and online forms) evolved during the implementation process of the catalog, a process that included several small modifications and one drastic change concerning the conceptual model. The latter resulted from improving the representation of different versions of the curriculum, due to changes in the local examination regulations (ER). We had to learn that, in different ER versions, existing LOs are included in different modules and whole modules may belong to different terms. Thus, some of the relations given in the core data model (Figure 2) needed to become dependent on the ER versions. Technically, this was achieved by introducing so-called semantic internal objects. The flexible way of imposing semantic constraints by semantic forms and templates (Figure 1) greatly supported schema versioning and content migration.

Additionally, for organizational reasons, a migration of ACLO-Web to different server hardware was necessary twice. In spite of an ongoing acquisition process, the system evolution went remarkably smoothly.

Views and Queries Based on the Core Data Model

The online LO catalog provides different ways of viewing and retrieving LOs. In general, there are 2 methods of access: (1) browsing predefined overviews, and (2) searching by criteria individually defined by the user. The model introduced above grants 3 independent views: users can select and view LOs

according to their position in the curricular context, their thematic focus, and the faculty and unit responsible for them, respectively.

Figure 5 shows the main page of the catalog. In the left-hand area, the page offers access via overviews, and in the right-hand area, the user can access search forms. All overviews are implemented as predefined searches (ie, overviews are not created manually, but instead are generated by the system using inline queries). Thus, all overviews are based on the actual state of the LO repository, and immediately show changes concerning added, modified, or deleted LO information.

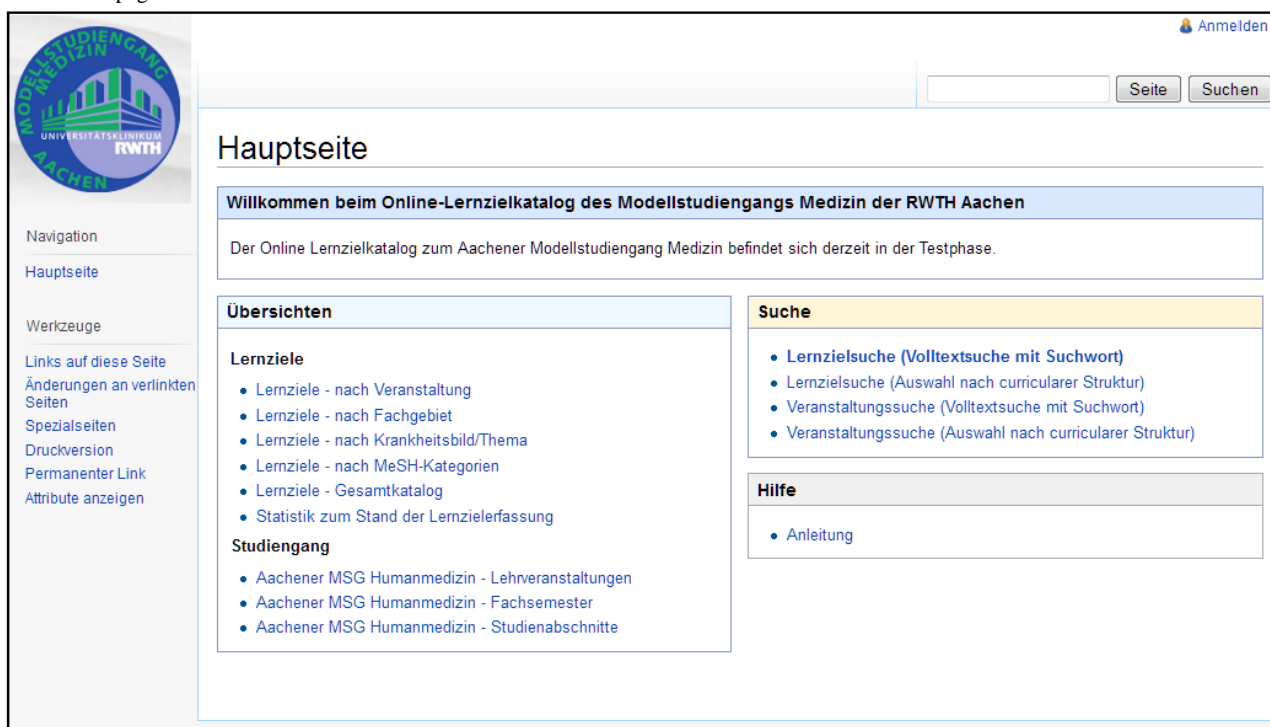
In order to use a predefined overview, students and faculty members first choose a curricular module, a teaching unit, a medical topic, or a MeSH term. The system then shows an overview of the LOs associated with this entity in a tabular layout. Detailed information on a particular LO is accessible by following the Web link to the wiki page describing the LO (provided by the overview). All tables can be dynamically rearranged by clicking on the heading of one or more columns, indicating the sorting criterion. The default order presented first depends on the type of overview (Table 1).

The overviews based on medical topic and medical specialty provide a longitudinal projection, that is, the users can see at which point in the course of study the topic is addressed or the specialty is involved in the curriculum, respectively.

The forms for searching by criteria enable (1) the retrieval of LOs by selecting the respective building blocks/segments of the curriculum, ie, module, term, part of the study, or examination regulation, and (2) a full-text search (supported by the auto-completion of terms used in the description of LOs).

Table 1. Default ordering corresponding to different types of LO overviews.

Overview of concerns	Sorting categories
Curricular module	Order in which LOs are addressed in the module
Medical specialty responsible for LO	Module associated with the LO (following the course of the curriculum)
Medical topic	Module associated with the LO (following the course of the curriculum)
MeSH based LO category	Short title of the LO (in lexicographic order)

Figure 5. Main page of ACLO-Web.

Collaborative Specification of Learning Objectives

As a key feature of the Social Semantic Web-based approach, the online catalog allows a distributed, collaborative, and structured data entry, requiring the client to have nothing more than a Web browser (and of course an Internet connection). The wiki-based user interface uses 13 online forms for data entry and 4 online forms for flexible LO search.

Figure 6 shows a screenshot of the form for describing LOs. Due to the high number of input widgets and the German labels of the system, the screenshot comes with additional English labels. The attributes used for specifying an LO have been introduced in the methods section above. In order to support offline data collection, we equipped the representatives responsible for the LO specification with predefined Excel forms that they could alternatively use for specifying LOs. The LO specifications were afterwards transferred to the system by the LO catalog team, and then checked online by the representatives.

Following the elicitation process described above, primary data entry was centralized and carried out by the ACLO team in

order to foster high and uniform data quality, and consistent semantic indexing. The faculty was then encouraged to revise the catalog. Following the wiki approach, ACLO-Web avoids the fine, granular access restrictions that are present in many content management systems. Thus, versioning support and page comments offered by the platform play a pivotal role in supporting an open, but nonetheless socially controlled, revision process.

Global Consistency Checking: ACLO-CM

Figure 7 shows a screenshot of the component for checking global consistency criteria (ACLO-CM). At present, ACLO-CM defines and checks 12 OCL constraints. Additionally, ACLO provides a convenient overview of modified SMW entries, which is structured by the categories of the model (see Figure 1). The report on existing consistency problems generated by ACLO-CM is structured by the categories of the affected SMW entries and the criteria checked. Figure 7 shows a typical ACLO-CM report.

Figure 6. Form for defining a medical learning objective.

The form is titled "Erstelle EingabeLernzielMedizin". It contains several sections:

- Category/Title:** "Kategorie* / Lernziel: Titel:" with a dropdown menu.
- MeSH-Schlüsselwort:** A dropdown menu.
- Geht zu Thema*:** A dropdown menu.
- Geht zu Fachgebiet*:** A dropdown menu.
- Geht zu Veranstaltung*:** A dropdown menu.
- Standard-Teilformulierung*:** A dropdown menu.
- Beschreibung*:** A large text area.
- Kompetenzniveau (kognitiv):** Radio buttons for "benennen", "erklären", "erkennen".
- Kompetenzniveau (psychomotorisch):** Radio buttons for "demonstrieren", "anwenden", "erfragen", "durchführen".
- Kompetenzniveau (affektiv):** Radio buttons for "einschätzen", "bewerten".
- Zusatzinformationen:** A text area.
- Literatur:** A text area.
- Weblinks (URL):** A text area.
- Hat Lernziel als Voraussetzung:** A dropdown menu.
- Hat empfohlenes Prüfungsformat:** Radio buttons for "MC-Klausur", "Klausur", "OSPE", "OSCE", "Mündlich".
- Vorgesehene Lernform:** Radio buttons for "Vorlesung", "Seminar", "Praktikum", "Eigenstudium".
- IMPP:** A text area.
- Schlagwörter:** A text area.
- Zugehörige ICD10 Codes auswählen:** A dropdown menu with a "Weitere hinzufügen" button.

Labels on the right side of the form point to the following fields:

- LO category
- Title
- Topic
- Med. Speciality
- Module
- Intro
- Description
- Competence Level
- Additional information (URL, Literature)
- Prerequisites
- Recommended form of assessment
- Recommended teaching format
- IMPP-Code
- ICD10-Codes

Figure 7. Report on global consistency problems generated by the ACLO-CM component.

The report is titled "Konsistenzprüfung der im Wiki angelegten Curricula" and is for "Beispiel - DA2011". It shows several consistency problems:

- Überprüfung der zugehörigen Studienabschnitte:**
 - Problem: Es gibt mehrere Studienabschnitte ohne (existierende) Vorgänger!
 - Beispiel - DA2011 - StAbs 01 - Einführung: Kein Vorgänger angegeben
 - Beispiel - DA2011 - StAbs 01 - Einführung Fehler: Kein Vorgänger angegeben
 - Problem: Auf den Studienabschnitt 'Beispiel - DA2011 - StAbs 02 - Fortgeschritten' folgen 2 Studienabschnitte!
 - Beispiel - DA2011 - StAbs 03 - Profi: folgt auf Beispiel - DA2011 - StAbs 02 - Fortgeschritten
 - Beispiel - DA2011 - StAbs 01 - SemiProfi: folgt auf Beispiel - DA2011 - StAbs 02 - Fortgeschritten
- Überprüfung der zugehörigen Fachsemester:**
 - Problem: Das Fachsemester 'Beispiel - DA2011 - StAbs 01 - SemiProfi - Fachsemester 06' dürfte nicht auf das Fachsemester 'Beispiel - DA2011 - StAbs 03 - Profi - Fachsemester 05' folgen, da dieses weder im selben noch im vorherigen Studienabschnitt liegt!
 - Problem: Auf das Fachsemester 'Beispiel - DA2011 - StAbs 02 - Fortgeschritten - Fachsemester 03' folgen 2 Fachsemester!
 - Beispiel - DA2011 - StAbs 02 - Fortgeschritten - Fachsemester 04: folgt auf Beispiel - DA2011 - StAbs 02 - Fortgeschritten - Fachsemester 03
 - Beispiel - DA2011 - StAbs 01 - Einführung - Fachsemester 07: folgt auf Beispiel - DA2011 - StAbs 02 - Fortgeschritten - Fachsemester 03

Formative Usability Study

The formative usability study was carried out in August 2012. Five medical students (in their second and third years of study) volunteered to participate in the study. Figure 8 shows the results of the overall rating of usability items after completing all 7 scenarios. The qualitative analysis of the full-text answers

yielded 30 separate statements assigned to 17 codes by the analysis.

Table 2 gives an overview of the code categories, codes, and numbers of statements assigned. Particular usability problems were raised in 10 statements, which addressed problems with the navigation and orientation (3 statements), comprehensibility, especially the interpretation of the short title of the LO (3

statements) (4 statements in total), and the presentation of the LO description (3 statements). These last statements said that the LO presentation should avoid the repeated presentation of the standard introductory text and should integrate the competence levels (verbs) into the textual description of the behavior.

With respect to the actual state of the LO repository, missing LOs were reported; some participants named missing aspects in detail (5 statements). The participants found very similar LOs repeatedly addressed by different medical specialties (4 statements), were not able to find a suitable starting point for their search in the case of scenario S5 (2 statements),

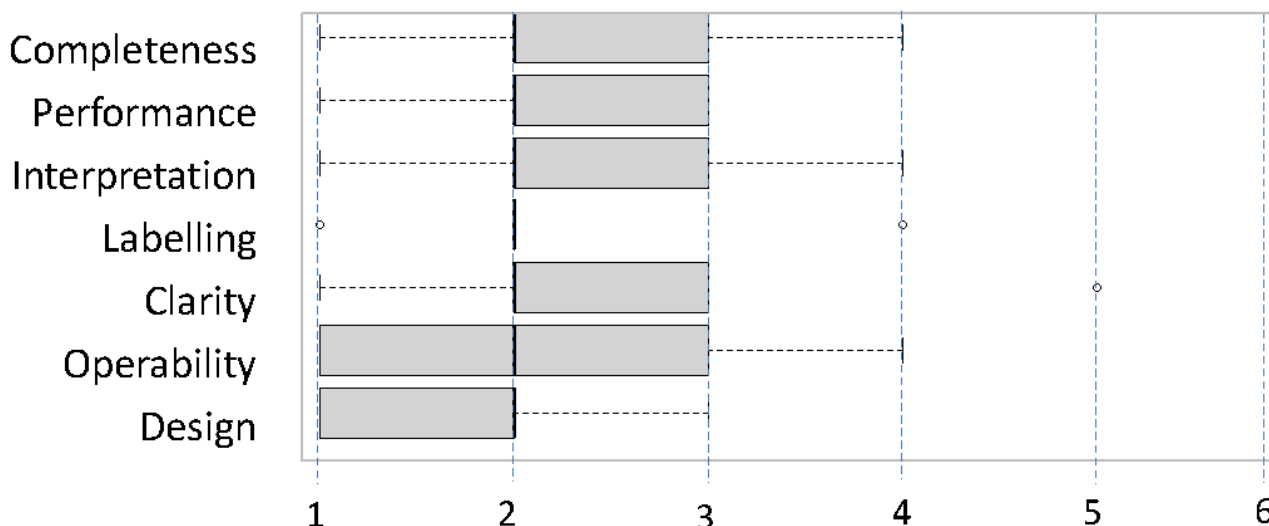
and—concerning the same issue—suggested the introduction of an attribute indicating multidisciplinary teaching formats (2 statements).

The statements contained 5 detailed recommendations, including the one proposing the attribute for multidisciplinary teaching formats. The extended search forms were recommended (2 statements); one participant suggested a graphical timeline for visualizing the longitudinal distribution of LOs.

The participants explicitly encouraged the alternative ways of accessing the LOs by overviews (1 statement), the structured, table-based presentation of the results (2 statements), and the longitudinal projection of a topic in the curriculum.

Table 2. Qualitative feedback: overview of the code categories derived from the statements.

Code category	Code	# of statements
Navigation/Orientation	Confusing SMW links	2
	Navigation/orientation problem	1
Comprehensibility	Interpretation of LO short title	3
	Interpretation of MeSH category	1
LO presentation	Competence levels as text	1
	Redundant introductory texts	2
State of the catalog	Missing LO	5
	Similar LO addressed by different specialty	4
Insufficient support for search	Unclear starting point of the search	2
	No index for multidisciplinary modules	2
Recommendations	Extended search	2
	Timeline view	1
Encouragement	Alternative views/ways of access	1
	Longitudinal view	1
	Structured, table-based overview	2

Figure 8. Overall rating of usability aspects (Rating scale: 1 = “Very good” to 6 = “Unsatisfactory”).

Discussion

Principal Findings

The focus of this paper is to report on the application of Social Semantic Web technology to support the knowledge management process and to enable the implementation and collaborative maintenance of an online catalog of learning objectives of a complete medical curriculum. Thus, we address neither the issue of the effect of competence-based learning objectives on the learning outcome nor a summative evaluation of the quality of ACLO.

Nonetheless, we report the results of the formative usability study because they were an essential part of the implementation process. Although it is known that usability studies can successfully rely on a relatively small number of participants [36], in our case the number is at the low end even for a formative study. The selection of the participants may be further biased, because we asked the official medical student representatives to name students who would possibly volunteer to participate in the study. Thus, the participating students can be assumed to take more interest than average in the further development and improvement of the existing model curriculum. The study did not involve the faculty. Although a medical expert was part of the LO team, his feedback was continuously incorporated into the implementation process. Nonetheless, the usability study will have to be extended to faculty members in the future.

Implications of the Usability Study

Taking into account the limitations mentioned above, the usability study nonetheless produced extremely valuable feedback that led to the improvement of the online catalog. The participants gave detailed hints on problems concerning orientation within the catalog, comprehensibility, and performance of the system. The hints on navigation problems and the interpretation of the short titles have already led to improvements in the present version of the system. Confidence in the results of the study is increased by observing a marked accordance of quantitative and qualitative feedback (triangulation); the overall rating of usability aspects yields

good results (especially with respect to operability and design), while some participants (voluntarily) encouraged the approach (especially complimenting the flexible, structured overviews). The main criticism concerned not the application but the state of the catalog, which was indeed poor when the formative evaluation took place and has markedly improved since. The box plots yielded a low rating for the completeness of information. In the context of the same scenario, the qualitative feedback stated that information on multidisciplinary modules was missing, or proposed an additional wiki attribute for capturing the relevant information.

As a main implication, the study showed that the online catalog actually works as an instrument for quality improvement and communication concerning LOs; while looking for different LOs assigned to the same medical topic, the students detected very similar LOs that were specified independently by different specialties. These require further coordination and reconciliation. The authors did not know about these findings during the preparation of the scenarios.

Enabling Nonproprietary Solutions

As noted in the introduction, custom-built software dominates the field of curriculum management systems. The SMW-based approach fills the gap between individual and flexible requirements and existing software platforms for supporting curriculum management. The fact that a first version of ACLO was ready for production use only a few weeks after the start of the project indicates that the SMW platform enabled the developers to easily configure a system that fulfills the specific requirements of a given CM project and supports the needs of different users. Furthermore, the data model and the information content of ACLO are accessible via standardized interfaces (eg, SPARQL queries) and can be exported to the W3C-standardized OWL format. Both aspects foster semantic interoperability and cross-platform migration. From a technical perspective, the availability of the MediaWiki Web application programming interface (Web-API)—well known from Wikipedia—seems even more important. The Web-API requires nothing but a Web connection in order to enable other systems to interact with ACLO.

The available formats and interfaces of the SMW-based approach not only improve standardized information access and system interoperability, but also allow importing existing classifications or ontologies to the system. This feature enabled, for example, the import of the ICD into ACLO. Furthermore, the system's information content can be easily enriched by seamlessly including interwiki links to other wiki-based information sources including Wikipedia.

Overall, the SMW-based approach shows potential to provide a versatile, but partly standardized platform for curriculum management, especially supporting collaborative aspects—far from being “yet another IT tool”.

Comparison With Prior Work

As stated in the introduction, ontologies serve as a means for semantic indexing [21-24]. Semantic indexing in ACLO uses not only medical standard classification, but also a folksonomy—following the collaborative approach. If compared with existing second-generation curriculum management systems such as eMed [18] and CMap [20], ACLO does not yet offer the complete functionality found there; neither the management of learning portfolios nor the support of all administrative planning tasks (eg, the management of timetables or resources) are presently supported. In contrast, none of these systems profits from the semantic standards and collaborative approach of the Social Semantic Web.

The Semantic Web has long been discussed as a means for supporting curriculum development by representing the learning design and content of a curriculum [37]. Following this line of tradition, Tang and Rahman developed a system design, which combines an ontology of the domain covered by a curriculum (here, computing domain) with a wiki-based information system [38], and Segeninac et al proposed to apply Semantic Web technology to curriculum development by modeling metadata for learning opportunities [39]. Recently, Coccoli et al outlined the potential of semantic wikis for collaborative curriculum development [40]. None of these system concepts or demonstrators made its way to productive use and none was intended to support curriculum management in the field of medical education. To the best of the authors' knowledge, ACLO is pioneering as a medical curriculum management system rigorously based on a Social Semantic Web approach and platform.

Agile Knowledge Management by SMW

The growth of the number of LOs during the past year shows that the system successfully supports the elicitation and systematic collection of LOs. As shown in Figure 4, the environment successfully supports an active process of the elicitation of learning objectives. The rapid increase of the number of LOs during the past months can be explained by the preparation and announcement of a faculty-wide evaluation process; this has obviously operated as an incentive. We received positive feedback concerning the applicability and usability of the system, not only from participants in the usability study but also from the LO team and faculty involved. This may be an effect of the familiarity of nearly all users with Wikipedia;

the use of the MediaWiki platform seems to have led to a low threshold for using the LO catalog.

At present, ACLO uses the few user roles originally provided by the MediaWiki platform. Thus, all members of the faculty can edit the LOs contained in ACLO during the revision process. This fact is causing controversy—part of the faculty demands the introduction of fine-grained, role-based access control, while others advocate the existing approach, relying on versioning support, the transparency of the page history, and shared social media etiquette.

Kiessling et al report on a systematic consensus process that improved the definition of (outcome-based) learning objectives, which the authors consider an intrinsically collaborative process [41]. Their statement is in line with the finding that curriculum management is resource-intensive and requires systematic change-management [9]. Additionally, Wong and Roberts argued that the procedural nature of curriculum mapping requires ongoing IT-enabled feedback [2]. Wikis proved to enable complex consensus processes and collaborative planning in the medical domain [42]. We consider that the SSW-based (Web 3.0) approach is an ideal platform for enabling and efficiently supporting curriculum mapping. The SMW platform, its versioning support, and standardized import/export formats effectively enabled successful evolution of the system. The weak schema constraints imposed by the approach (Figure 2) allowed a very flexible evolution of the model, while at the same time enabling structured information management and access.

Thus, the SMW-based approach proved to enable an agile implementation of computer-supported knowledge management. The approach, based on standard Social Semantic Web formats and technology, represents a feasible and effectively applicable compromise between answering to the individual requirements of curriculum management at a particular medical school and using proprietary systems. Given the overall feasibility of a Web 3.0-based curriculum management system, these special aspects of a flexible and agile knowledge management can indeed foster collaborative curriculum mapping as a feedback-driven process.

Future Directions

As mentioned before, further formative usability testing and the scheduled revision phase will involve all members of the faculty. ACLO will be open for all students at our medical school in approximately April 2013. We are now preparing a summative evaluation of the system, including a log file analysis of the users' behavior. At present, 2 questionnaires (addressing students and faculty, respectively) will undergo pilot testing and a revision process. The questionnaire for faculty contains scaled items concerning qualitative aspects of the learning objectives' specifications (SMART-criteria: Specific, Measurable, Accepted, Realistic, Timely), the completeness, and the perceived effectiveness of the catalog. The students' questionnaire further addresses the students' appraisal of the benefit from ACLO. A second curriculum (dentistry) is going to be included in 2014.

Maloney et al showed that medical students appreciate the benefit of online repositories of learning resources [43]. They also showed that milestones of the curriculum (eg, examinations) often trigger access to the repositories. Consequently, ACLO will be integrated with an already existing system used for the consistent indexing of eLearning media, by a semantic network based on the Medical Subject Headings and ICD codes, which were linked by associations taken from the SNOMED Clinical Terms terminology. The integration will link the topics and

external vocabularies associated with ACLO learning objectives to concepts of the semantic network and will therefore extend the systems capability of finding similar LOs.

Last but not least, the ongoing project of a German national catalog of competence-based medical learning objectives will heavily influence the future development of ACLO-Web and result in challenging tasks concerning system integration and LO identification. This will eventually be based on ontology mapping approaches.

Acknowledgments

The authors would like to thank all faculty of the RWTH Aachen Medical School involved in the implementation of ACLO and the anonymous reviewers, who helped to improve the paper substantially.

Conflicts of Interest

None declared.

References

1. Denny JC, Smithers JD, Armstrong B, Spickard A. "Where do we teach what?" Finding broad concepts in the medical school curriculum. *J Gen Intern Med* 2005 Oct;20(10):943-946 [FREE Full text] [doi: [10.1111/j.1525-1497.2005.0203.x](https://doi.org/10.1111/j.1525-1497.2005.0203.x)] [Medline: [16191143](https://pubmed.ncbi.nlm.nih.gov/16191143/)]
2. Wong RY, Roberts JM. Real time curriculum map for internal medicine residency. *BMC Med Educ* 2007;7:42 [FREE Full text] [doi: [10.1186/1472-6920-7-42](https://doi.org/10.1186/1472-6920-7-42)] [Medline: [17988402](https://pubmed.ncbi.nlm.nih.gov/17988402/)]
3. Willett TG. Current status of curriculum mapping in Canada and the UK. *Med Educ* 2008 Aug;42(8):786-793. [doi: [10.1111/j.1365-2923.2008.03093.x](https://doi.org/10.1111/j.1365-2923.2008.03093.x)] [Medline: [18564300](https://pubmed.ncbi.nlm.nih.gov/18564300/)]
4. Nikendei C, Weyrich P, Jünger J, Schrauth M. Medical education in Germany. *Med Teach* 2009 Jul;31(7):591-600. [Medline: [19811144](https://pubmed.ncbi.nlm.nih.gov/19811144/)]
5. Hausman JJ. Mapping as an approach to curriculum planning. *Curriculum Theory Network* 1974;4(2/3):192-198.
6. English FW. Curriculum Mapping. *Educational Leadership* 1980;37(7):558-559.
7. Gjerde CL. 'Curriculum mapping': objectives, instruction, and evaluation. *J Med Educ* 1981 Apr;56(4):316-323. [Medline: [7218296](https://pubmed.ncbi.nlm.nih.gov/7218296/)]
8. Harden RM. AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Med Teach* 2001 Mar;23(2):123-137. [doi: [10.1080/01421590120036547](https://doi.org/10.1080/01421590120036547)] [Medline: [11371288](https://pubmed.ncbi.nlm.nih.gov/11371288/)]
9. Fleiszer DM, Posel NH. Development of an undergraduate medical curriculum: the McGill experience. *Acad Med* 2003 Mar;78(3):265-269. [Medline: [12634204](https://pubmed.ncbi.nlm.nih.gov/12634204/)]
10. Lee MY, Albright SA, Alkasab T, Damassa DA, Wang PJ, Eaton EK. Tufts Health Sciences Database: lessons, issues, and opportunities. *Acad Med* 2003 Mar;78(3):254-264. [Medline: [12634202](https://pubmed.ncbi.nlm.nih.gov/12634202/)]
11. Salas AA, Anderson MB, LaCourse L, Allen R, Candler CS, Cameron T, et al. CurrMIT: a tool for managing medical school curricula. *Acad Med* 2003 Mar;78(3):275-279. [Medline: [12634208](https://pubmed.ncbi.nlm.nih.gov/12634208/)]
12. Landry LG, Alameida MD, Orsolini-Hain L, Renwanz Boyle A, Privé A, Chien A, et al. Responding to demands to change nursing education: use of curriculum mapping to assess curricular content. *J Nurs Educ* 2011 Oct;50(10):587-590. [doi: [10.3928/01484834-20110630-02](https://doi.org/10.3928/01484834-20110630-02)] [Medline: [21710963](https://pubmed.ncbi.nlm.nih.gov/21710963/)]
13. Bloch R, Bürgi H. The Swiss catalogue of learning objectives. *Med Teach* 2002 Mar;24(2):144-150. [doi: [10.1080/01421590220120759](https://doi.org/10.1080/01421590220120759)] [Medline: [12098433](https://pubmed.ncbi.nlm.nih.gov/12098433/)]
14. Bosse HM, Dambe R, Juenger J, Kadmon M. An interdisciplinary and interactive online tool to manage the continuous development of learning objectives in a curriculum. *Z Evid Fortbild Qual Gesundheitswes* 2011;105(2):116-123. [doi: [10.1016/j.zefq.2010.11.001](https://doi.org/10.1016/j.zefq.2010.11.001)] [Medline: [21496780](https://pubmed.ncbi.nlm.nih.gov/21496780/)]
15. Hege I, Siebeck M, Fischer MR. An online learning objectives database to map a curriculum. *Med Educ* 2007 Nov;41(11):1095-1096. [doi: [10.1111/j.1365-2923.2007.02873.x](https://doi.org/10.1111/j.1365-2923.2007.02873.x)] [Medline: [17883386](https://pubmed.ncbi.nlm.nih.gov/17883386/)]
16. Hege I, Nowak D, Kolb S, Fischer MR, Radon K. Developing and analysing a curriculum map in Occupational- and Environmental Medicine. *BMC Med Educ* 2010;10:60 [FREE Full text] [doi: [10.1186/1472-6920-10-60](https://doi.org/10.1186/1472-6920-10-60)] [Medline: [20840737](https://pubmed.ncbi.nlm.nih.gov/20840737/)]
17. Dreinhöfer KE, Walcher F, Obertacke U, Waydhas C, Josten C, Rüsseler M, et al. [Development of a catalogue of undergraduate learning objectives for orthopaedics and traumatology]. *Z Orthop Unfall* 2008;146(4):520-533. [doi: [10.1055/s-2008-1038899](https://doi.org/10.1055/s-2008-1038899)] [Medline: [18704850](https://pubmed.ncbi.nlm.nih.gov/18704850/)]

18. Watson EG, Moloney PJ, Toohey SM, Hughes CS, Mobbs SL, Leeper JB, et al. Development of eMed: a comprehensive, modular curriculum-management system. *Acad Med* 2007 Apr;82(4):351-360. [doi: [10.1097/ACM.0b013e3180334d41](https://doi.org/10.1097/ACM.0b013e3180334d41)] [Medline: [17414190](https://pubmed.ncbi.nlm.nih.gov/17414190/)]
19. Bell CE, Ellaway RH, Rhind SM. Getting started with curriculum mapping in a veterinary degree program. *J Vet Med Educ* 2009;36(1):100-106. [doi: [10.3138/jvme.36.1.100](https://doi.org/10.3138/jvme.36.1.100)] [Medline: [19435996](https://pubmed.ncbi.nlm.nih.gov/19435996/)]
20. Oyewumi M, Isaac K, Schreiber M, Campisi P. Undergraduate otolaryngology education at the University of Toronto: a review using a curriculum mapping system. *J Otolaryngol Head Neck Surg* 2012 Feb;41(1):71-75. [Medline: [22498272](https://pubmed.ncbi.nlm.nih.gov/22498272/)]
21. Denny JC, Smithers JD, Miller RA, Spickard A. "Understanding" medical school curriculum content using KnowledgeMap. *J Am Med Inform Assoc* 2003;10(4):351-362 [FREE Full text] [doi: [10.1197/jamia.M1176](https://doi.org/10.1197/jamia.M1176)] [Medline: [12668688](https://pubmed.ncbi.nlm.nih.gov/12668688/)]
22. Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. *J Am Med Inform Assoc* 2010 Jun;17(3):229-236. [doi: [10.1136/jamia.2009.002733](https://doi.org/10.1136/jamia.2009.002733)] [Medline: [20442139](https://pubmed.ncbi.nlm.nih.gov/20442139/)]
23. Willett TG, Marshall KC, Broudo M, Clarke M. TIME as a generic index for outcome-based medical education. *Med Teach* 2007 Sep;29(7):655-659. [doi: [10.1080/01421590701615808](https://doi.org/10.1080/01421590701615808)] [Medline: [18236252](https://pubmed.ncbi.nlm.nih.gov/18236252/)]
24. Dexter J, Koshland G, Waer A, Anderson D. Mapping a curriculum database to the USMLE Step 1 content outline. *Med Teach* 2012;34(10):e666-e675. [doi: [10.3109/0142159X.2012.687477](https://doi.org/10.3109/0142159X.2012.687477)] [Medline: [23088357](https://pubmed.ncbi.nlm.nih.gov/23088357/)]
25. Gasevic D, Zouaq A, Torniai C, Jovanovic J, Hatala M. An Approach to Folksonomy-Based Ontology Maintenance for Learning Environments. *IEEE Trans Learning Technol* 2011 Oct;4(4):301-314. [doi: [10.1109/TLT.2011.21](https://doi.org/10.1109/TLT.2011.21)]
26. Bruner J. *The process of education*. Cambridge: Harvard University Press; 1977.
27. Krötzsch M, Vrandečić D, Völkel M. Semantic MediaWiki. In: Cruz I, Decker S, Allemang D, Preist C, Schwabe D, Mika P, editors. *The Semantic Web - ISWC 2006*. Heidelberg: Springer; 2006:935-942.
28. Berners-Lee T, Hendler J. Publishing on the semantic web. *Nature* 2001 Apr 26;410(6832):1023-1024. [doi: [10.1038/35074206](https://doi.org/10.1038/35074206)] [Medline: [11323639](https://pubmed.ncbi.nlm.nih.gov/11323639/)]
29. Horrocks I, Patel-Schneider PF, Van Harmelen F. From SHIQ and RDF to OWL: The making of a web ontology language. *Web semantics: science, services and agents on the World Wide Web* 2003;1(1):7-26.
30. McGuinness DL, Van Harmelen F. OWL web ontology language overview. *W3C recommendation* 2004;10(2004-03):1 [FREE Full text]
31. Prud'hommeaux E, Seaborne A. SPARQL query language for RDF. *W3C recommendation* 2008;15:1 [FREE Full text]
32. The LimeSurvey project team. LimeSurvey Homepage. URL: <http://www.limesurvey.org/> [accessed 2013-03-14] [WebCite Cache ID 6F73evx9I]
33. The R Foundation for Statistical Computing. The R Project for Statistical Computing. URL: <http://www.r-project.org/> [accessed 2013-03-14] [WebCite Cache ID 6F73oyxZ8]
34. Seidler KO, Vogelgesang K. Apache Friends Homepage. URL: <http://www.apachefriends.org/en/index.html> [accessed 2013-03-14] [WebCite Cache ID 6F7B8WavH]
35. Wikimedia Foundation. MediaWiki Homepage. URL: <http://www.mediawiki.org/wiki/MediaWiki> [accessed 2013-03-14] [WebCite Cache ID 6F7BIR8ek]
36. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Ashlund S, Mullet K, Henderson A, Hollnagel E, White T, editors. *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. New York: ACM Press; 1993:206-213.
37. Koper R. Use of the Semantic Web to Solve Some Basic Problems in Education. *Journal of Interactive Media in Education* 2004;6:1-23 [FREE Full text]
38. Tang A, Rahman AA. Proof-of-Concept Design of an Ontology-Based Computing Curricula Management System. In: Cherifi H, Zain JM, El-Qawasmeh E, editors. *Digital Information and Communication Technology and Its Applications*. Heidelberg: Springer; 2011:280-292.
39. Segedinac M, Konjovic Z, Surla D, Savic G. An OWL representation of the MLO model. In: *Proceedings of the IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics (SISY)*. 2012 Presented at: 2012 IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics (SISY); Sept. 20-22, 2012; Serbia p. 465-470.
40. Coccoli M, Vercelli G, Vivinet G. Semantic Wiki: a collaborative tool for instructional content design. *Journal of e-Learning and Knowledge Society* 2012;8(2):113-122.
41. Kiessling C, Dieterich A, Fabry G, Hölzer H, Langewitz W, Mühlinghaus I, Committee Communication Social Competencies of the Association for Medical Education Gesellschaft für Medizinische Ausbildung, Basel Workshop Participants. Communication and social competencies in medical education in German-speaking countries: the Basel consensus statement. Results of a Delphi survey. *Patient Educ Couns* 2010 Nov;81(2):259-266. [doi: [10.1016/j.pec.2010.01.017](https://doi.org/10.1016/j.pec.2010.01.017)] [Medline: [20223614](https://pubmed.ncbi.nlm.nih.gov/20223614/)]
42. Gupta S, Wan FT, Newton D, Bhattacharyya OK, Chignell MH, Straus SE. WikiBuild: a new online collaboration process for multistakeholder tool development and consensus building. *J Med Internet Res* 2011;13(4):e108 [FREE Full text] [doi: [10.2196/jmir.1833](https://doi.org/10.2196/jmir.1833)] [Medline: [22155694](https://pubmed.ncbi.nlm.nih.gov/22155694/)]
43. Maloney S, Chamberlain M, Morrison S, Kotsanas G, Keating JL, Ilic D. Health professional learner attitudes and use of digital learning resources. *J Med Internet Res* 2013;15(1):e7 [FREE Full text] [doi: [10.2196/jmir.2094](https://doi.org/10.2196/jmir.2094)] [Medline: [23324800](https://pubmed.ncbi.nlm.nih.gov/23324800/)]

Abbreviations

ACLO: Aachen Catalogue of Learning Objectives

ACLO-CM: ACLO consistency module

ACLO-Web: Web-based version of the Aachen Catalogue of Learning Objectives

AMMC: Aachen Medical Model Curriculum

ER: examination regulations

HSDB: Health Science Database

ICD: International Classification of Diseases

IMPP: Institute for Medical and Pharmaceutical Examination Items

LO: learning objective

MeSH: Medical Subject Headings

OCL: Object Constraint Language

OPS: German modification of the International Classification of Procedure in Medicine (Operationen- und Prozedurenschlüssel)

OWL: Web Ontology Language

SCLO: Swiss Catalogue of Learning Objectives

SMW: Semantic MediWiki

SSW: Social Semantic Web

Edited by G Eysenbach; submitted 17.03.13; peer-reviewed by M Broom, T Mabotuwana, B Good; comments to author 23.04.13; revised version received 31.05.13; accepted 14.06.13; published 15.08.13.

Please cite as:

Spreckelsen C, Finsterer S, Cremer J, Schenkat H

Can Social Semantic Web Techniques Foster Collaborative Curriculum Mapping In Medicine?

J Med Internet Res 2013;15(8):e169

URL: <http://www.jmir.org/2013/8/e169/>

doi: [10.2196/jmir.2623](https://doi.org/10.2196/jmir.2623)

PMID: [23948519](https://pubmed.ncbi.nlm.nih.gov/23948519/)

©Cord Spreckelsen, Sonja Finsterer, Jan Cremer, Hennig Schenkat. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 15.08.2013. 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.

Viewpoint

Security Analysis and Improvements to the PsychoPass Method

Bostjan Brumen^{1*}, BSc, PhD; Marjan Heričko^{1*}, MSc, PhD; Ivan Rozman^{1*}, MSc, PhD; Marko Hölbl^{1*}, BSc, PhD

Institute of Informatics, Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia

* all authors contributed equally

Corresponding Author:

Marko Hölbl, BSc, PhD

Institute of Informatics

Faculty of Electrical Engineering and Computer Science

University of Maribor

Smetanova 17

Maribor, 2000

Slovenia

Phone: 386 2 2207292

Fax: 386 2 2207272

Email: marko.holbl@uni-mb.si

Abstract

Background: In a recent paper, Pietro Cipresso et al proposed the PsychoPass method, a simple way to create strong passwords that are easy to remember. However, the method has some security issues that need to be addressed.

Objective: To perform a security analysis on the PsychoPass method and outline the limitations of and possible improvements to the method.

Methods: We used the brute force analysis and dictionary attack analysis of the PsychoPass method to outline its weaknesses.

Results: The first issue with the Psychopass method is that it requires the password reproduction on the same keyboard layout as was used to generate the password. The second issue is a security weakness: although the produced password is 24 characters long, the password is still weak. We elaborate on the weakness and propose a solution that produces strong passwords. The proposed version first requires the use of the SHIFT and ALT-GR keys in combination with other keys, and second, the keys need to be 1-2 distances apart.

Conclusions: The proposed improved PsychoPass method yields passwords that can be broken only in hundreds of years based on current computing powers. The proposed PsychoPass method requires 10 keys, as opposed to 20 keys in the original method, for comparable password strength.

(*J Med Internet Res* 2013;15(8):e161) doi:[10.2196/jmir.2366](https://doi.org/10.2196/jmir.2366)

KEYWORDS

security; passwords; cryptanalysis

Introduction

In a recent paper, El Emam, Moreau, and Jonker highlighted the importance of using strong passwords to protect personal health information in clinical trials stored in files [1]. In their settings, we have a typical offline password guessing scheme.

Pietro Cipresso et al have commented on the paper by elaborating on the potential problem people may have creating passwords that are complex but at the same time easy to remember, and they propose a solution called the PsychoPass method [2]. The method is discussed in the context of user-protected files; however, it can also be used in other settings

requiring a password, including administrator account passwords.

The proposed solution by Cipresso et al has some limitations to be considered. But before we describe the limitations, we present a short discussion on strong passwords, as both papers have omitted an explanation of why the passwords can be weak and how serious the weaknesses can be.

First, the general rule is that the strength of passwords is proportional to their length and to the type of symbols being used, provided that the password's symbols are drawn randomly from a pool of possible symbols—the key concept here is randomness [3]. The formula expressing the number of possible

combinations is $s=b^{le}$, where s is the total number of combinations, b is the total number of symbols in a domain and le is the length of the password in number of characters. Typically, the domain where the password symbols are drawn from consists of lowercase (a...z) and uppercase letters (A...Z), numbers (0...9) and special symbols (eg, “#\$%&/). In English, the number of letters is 52 (26 lower case + 26 upper case), in addition to 10 numbers (0...9) and some symbols (eg, 13), totaling 75. If, for example, up to 7 characters are used for a password, there is a total of $s=75^1+75^2+75^3+\dots+75^7=75^8-1=1,001,129,150,390,625$ combinations. However, an adversary may try out all the combinations in a so-called brute force attack. Today, the reported speeds (v) are in the range of $v=10^9$ combinations/second [4-7], up to 10^{12} combinations/second [3], for restoring a plaintext password from a given hash value, as was the scenario in the paper by El Emam et al [1]. The same scenario—an attack using precalculated hash values for cracking passwords generated by the PsychoPass method—can be used.

To find out how much time it takes to check all the combinations (in a worst-case scenario where the sought-after password is the very last one to try), we use the well-known equation from high-school physics, $t=s/v$, where time (the object has traveled) is the distance divided by velocity, which is in our case the number of combinations divided by the speed of how many combinations could be tested in 1 second. The calculation is based upon an assumption that the hash values of all possible combinations of letters are precomputed: $t=s/v=1,001,129,150,390,625$ comb. / 1,000,000,000 comb. / second=1,001,129.15 seconds=16,685.49 minutes=278.09 hours=11.59 days. That is, it will take at most 11.59 days (11 days, 14 hours, 5 minutes, 29 seconds) to find the password. On average however, the time is halved (5 days, 19 hours, 2 minutes, 45 seconds).

Due to Moore’s Law [8] still in effect, the computing power doubles roughly every 18 months. In 10 years, we can expect the speeds to be 100 times faster than today’s speed. One should note, however, that the estimate of the increase of speeds based on Moore’s law is rather conservative. The techniques and algorithms advance faster than brute computing power (eg, using GPUs and server farms for hire [9]). Also, quantum computing will have an enormous impact on security [10].

Tables 1 and 2 list the number of all possible combinations for passwords of length up to 7 and 9, respectively, for different sizes of character pools (25=lower case letters only, 50=lower and upper case letters, 60=letters and numbers, 75=letters, numbers, and special symbols), the time required to check *all*

possible combinations at today’s speeds, and the time it will take in 10 years from now due to speed improvements.

The tables describe a well-known phenomenon in the information security field: what is safe today will most probably not be safe tomorrow [3]. Consider the setting described in the original paper by El Emam et al [1], where personal health information in clinical trials is stored in files. These same files will be around for years and although an attacker might not have enough computing power today, he or she will have it a few years from now. If the files contain interesting medical data about a celebrity or a public person, for example, the disclosure will be as damaging in the future as it might be today.

Second, the general rule—the more symbols in a password the better—has an important exception: if the password is otherwise long, but is a word, it is considered to be a weak password (the characters are no longer drawn randomly, but from a specific distribution). Such a password is susceptible to a dictionary attack [3]. Suppose an adversary composes a dictionary of all words of all languages and calculates the corresponding hash value. While it is hard to tell how many languages there are in the world—the estimates vary from 5000 to 10,000 languages—there are 6909 living human languages catalogued [11]. Let us suppose that each language contains 1 million words—a recent study [12] estimated the number of words in the English lexicon was 1,022,000 in 2000. Based on these numbers, we estimate that the total number of human words would not exceed 7,060,998,000 words; the actual number is much lower due to overlapping of words between languages. It would take very little time for an adversary to try all the words as passwords: $t=s/v=7,060,998,000$ words / 1,000,000,000 words /s=7.06 seconds.

Today, dictionaries containing words and passwords have several billion entries [13,14]. When choosing the correct length of a password, it is essential to observe one of the basic principles of security [3]: (1) a password scheme is said to be computationally secure if the cost of breaking the cipher exceeds the value of the protected information, or (2) the time required to break the password exceeds the useful lifetime of the information. Today’s costs for building a password-cracking machine are negligible [4,6,7], so the only principle to rely on is the time required for breaking the password. For the purpose of this paper, let us assume that the useful lifetime of the (medical) information is 60 years. Under this assumption, a safe password would be made of at least 9 characters from upper and lowercase letters, numbers, and symbols.

The aim of this paper is to examine the strength of the PsychoPass method in light of these assumptions.

Table 1. Number of combinations of passwords of length up to 7 and maximum time to crack them now and in 10 years.

Size	No. of combinations	Time to crack now	Time to crack in 10 years from now
25	1,52588E+11	2 minutes, 33 seconds	0 seconds
50	3,90625E+13	10 hours, 51 minutes, 2 seconds	39 seconds
60	1,67962E+14	1 day, 22 hours, 39 minutes, 22 seconds	2 minutes, 48 seconds
75	1,00113E+15	11 days, 14 hours, 5 minutes, 29 seconds	16 minutes, 41 seconds

Table 2. Number of combinations of passwords of length up to 9 and maximum time to crack them now and in 10 years.

Size	No. of combinations	Time to crack now	Time to crack in 10 years from now
25	9,53674E+13	1 day, 2 hours, 29 minutes, 27 seconds	1 minutes, 35 seconds
50	9,76563E+16	3 years, 35 days, 6 hours, 44 minutes, 10 seconds	1 day, 3 hours, 7 minutes, 36 seconds
60	6,04662E+17	19 years, 63 days, 9 hours, 36 minutes	6 days, 23 hours, 57 minutes, 42 seconds
75	5,63135E+18	178 years 207 days, 16 hours, 17 minutes, 51 seconds	65 days, 4 hours, 15 minutes, 51 seconds

Security Issues of the PsychoPass Method

We have identified two issues with the PsychoPass method proposed by Cipresso et al. First, their method works only on keyboards with the same layout. In many countries (eg, in Canada), there are several different keyboard layouts, rendering their method practically useless. For example, a sequence on a Canadian Multilingual Standard keyboard (see [Figure 1](#)) starting with key “w”, followed by combination “SHIFT” + key “3” produces password “w#”, while the same sequence repeated on a Canadian French keyboard ([Figure 2](#)) produces password “w/”.

The situation gets worse when the keyboards are of different types (QWERTZ vs QWERTY vs AWERTY) or when the method is used across various platforms, eg, from desktop PCs or desktops to mobile devices, such as Android-powered tablets (see layout in [Figure 3](#) vs [Figures 1](#) and [2](#)).

The idea of PsychoPass is that a password can be created, memorized, and recalled by just thinking of an action sequence instead of a word or string of characters [2]. When the keyboard layout is different, the user cannot reproduce the very same password as she or he only knows the sequence of the keys, but not the key values themselves.

The above-mentioned problem is merely technical and requires the user to use the same type of keyboard. Additionally, with some basic training or professional help, a user can change the keyboard layout without physically replacing the keyboard. With a different system layout, a user would again reproduce the correct password even on the physically different keyboard. The interoperability between traditional and mobile devices remains a minor challenge.

Second, and more importantly, the method proposed by Cipresso et al has a security design issue because it produces predictable passwords, being prone to brute force attack. The PsychoPass method, when implemented as demonstrated by the authors in their video (and as can be seen from the figures and described in the paper) [2], produces a password such that it starts at a certain key and then proceeds only to the first neighbor of that key, and then again only to the first neighbor of that key, until the sequence length is reached; the password sequence is then

repeated. Characters produced by such a procedure are not drawn randomly but are drawn according to some function (in this case, by the proximity function). As can be seen from the Cipresso et al’s [Figure 1](#), they have produced a sequence “f-t-6-t-y-g-r-5”, where each key in the sequence is a neighbor to the previous one.

Proposing the use of adjacent keys on a keyboard produces combinations that are not only nonrandom, but these combinations themselves form a dictionary of finite combinations. As the PsychoPass method is publicly disclosed, constructing an algorithm for a dictionary-based attack script for the PsychoPass method is exceedingly easy.

The total number of different combinations (s) using the demonstrated PsychoPass method is $s = n_k \cdot b^{le} \cdot n_s$, where n_k is the number of different characters on the keyboard from where the sequence can start, b is the number of possible next keys, le is the length of the produced sequence, and n_s is the number of repeated sequences.

At the beginning, we have some 45 key combinations ($n_k=45$) for selecting the key as the starting point (the authors chose key “f”). From there on, each keyboard key has (at most) 8 neighbors (plus the key itself), so in each step only 1 out of 9 combinations ($b=9$) can be used. The authors have created a 24-key password by repeating the same sequence of length 8 ($le=8$) three times ($n_s=3$) and have claimed that the password is a strong one. However, their claim is optimistic. The total number of different password combinations that can be produced by their method is $s = n_k \cdot b^{le} \cdot n_s = 45 \cdot 9^8 \cdot 3 = 5,811,307,355$. All such passwords can be checked in less than 6 seconds. [Table 3](#) lists the amount of time required to test all passwords created by the PsychoPass method for a different number of sequence repetitions and a different number of keystrokes in a sequence.

It can be observed that the number of repetitions of sequences (n_s) does not contribute significantly to the overall strength of the password. The only contributing factor is the number of letters in the password. For the proposed PsychoPass method to be considered safe and to produce strong password for today’s use, a user would have to remember a 17-key sequence. But when considering Moore’s law, a 20-key nonrepeating sequence should be used.

Table 3. Strength of original PsychoPass method for different parameter settings.

n _s	le=8	le=17	le=20
1	2 seconds	23 years, 291 days, 46 minutes, 16 seconds	173 years, 176 days, 4 hours, 10 minutes, 57 seconds
2	4 seconds	47 years, 217 days, 1 hour, 32 minutes, 33 seconds	346 years, 352 days, 8 hours, 21 minutes, 53 seconds
3	6 seconds	71 years, 143 days, 2 hours, 18 minutes, 49 seconds	520 years, 163 days, 12 hours, 32 minutes, 50 seconds
4	8 seconds	95 years, 69 days, 3 hours, 5 minutes, 6 seconds	693 years, 339 days, 16 hours, 43 minutes, 46 seconds
5	10 seconds	118 years, 360 days, 3 hours, 51 minutes, 22 seconds	867 years, 150 days, 20 hours, 54 minutes, 43 seconds
6	12 seconds	142 years, 286 days, 4 hours, 37 minutes, 39 seconds	1040 years, 327 days, 1 hours, 5 minutes, 39 seconds
7	14 seconds	166 years, 212 days, 5 hours, 23 minutes, 55 seconds	1214 years, 138 days, 5 hours, 16 minutes, 36 seconds
8	15 seconds	190 years, 138 days, 6 hours, 10 minutes, 12 seconds	1387 years, 314 days, 9 hours, 27 minutes, 33 seconds

Figure 1. Canadian multilingual standard keyboard layout.



Figure 2. Canadian French keyboard layout.

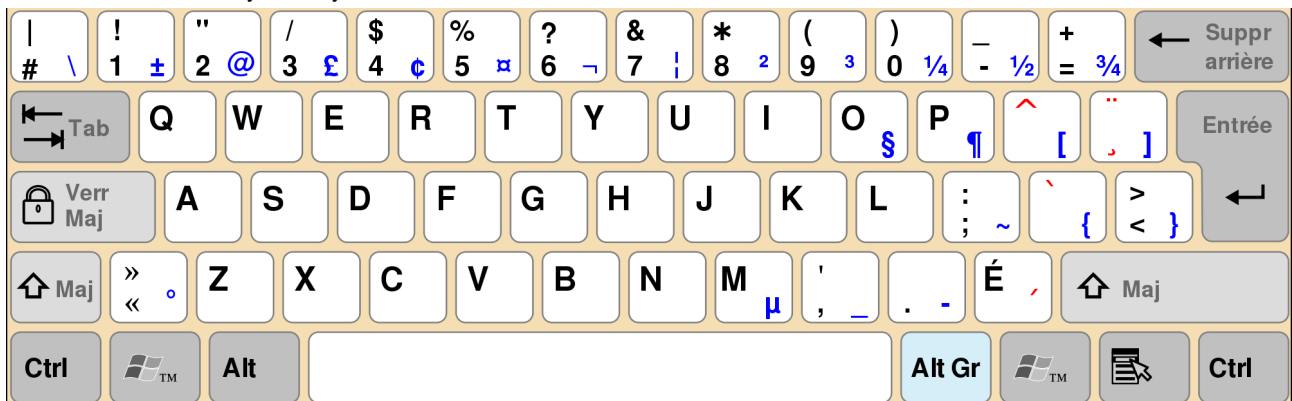
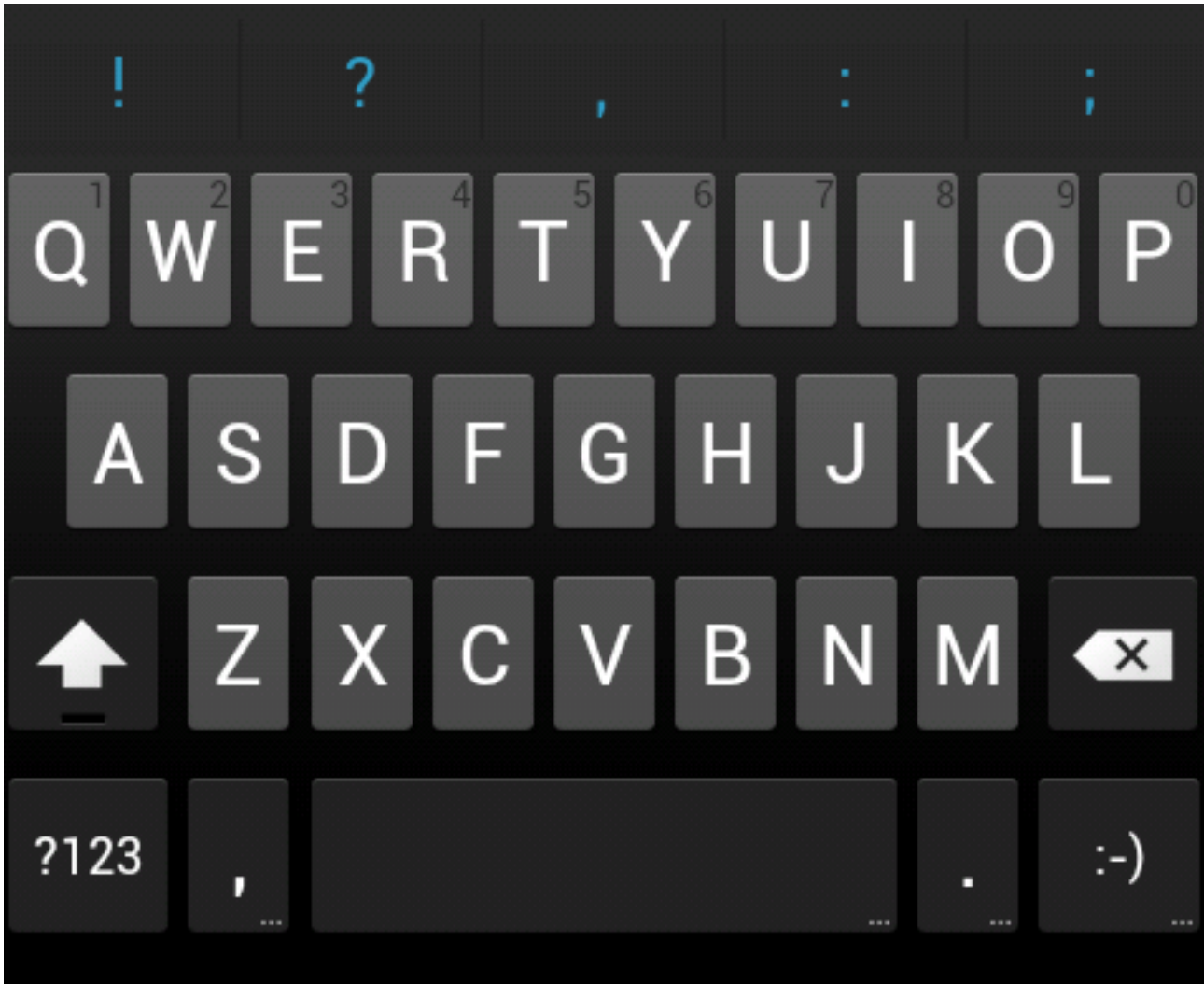


Figure 3. Android ICS keyboard layout.



Proposed Improvements to the PsychoPass Method

The PsychoPass method for generating passwords is based on a very interesting concept. The original version, which has a security issue, can be improved as follows. First, one should use the SHIFT key and ALT-GR key in combination with other keys. This way, the initial number of combinations (n_k) increases from 45 to some 110 (=45 characters without shift + 45 characters with shift + some 20 characters with ALT-GR).

Second, a user should remember the next key that is not only 1 distance away, but 1 or 2 (or more). For example, if “Q” is initially selected, then “E” (or “e”), should be taken into the account as well. Here, “E” has a distance of 2 units from “Q”. This way the number of possible next keys increases from 9 to

approximately 18 (approximately because it depends on the location of the key, eg, “B” has less two-unit distance neighbors than “E” since “B” is next to the space bar).

Now, combining the use of SHIFT and ALT-GR keys and the use of a larger distance between the keys increases the base (b) from 9 to approximately 54 (ie, 9 neighbors, each in combination with the plain key, SHIFT + key or ALT-GR + key). Combined, the total number of different passwords that can be produced by the improved method with 3 repetitions and sequence length of 9 characters is $s = n_k^{b \cdot l} \cdot n_s = 110 \cdot 54^9 \cdot 3 = 1,288,420,951,063,403,520$.

All these passwords can be checked in 40 years, 312 days, 6 hours, 42 minutes, and 31 seconds. This is a considerable improvement over the 6-second attack on the original method.

Table 4. Strength of improved PsychoPass method for different parameter settings (truncated to years).

n_s	1e=8	1e=9	1e=10
1	0	13	735
2	0	27	1470
3	0	40	2206

Table 4 lists the time required (in years) to check all the password combinations under different parameter settings. It can be observed (again) that the sequence length is the key contributing factor to the overall security of the password. For the proposed improvement to the PsychoPass method to be considered safe and to produce strong password for today's use, a user would have to remember a 10-key sequence, repeated only once.

Conclusion

The PsychoPass method, as proposed by Pietro Cipresso et al in [2], has two issues. The first issue is merely a technical one: the passwords can be produced and reproduced only on keyboards with the same keyboard layout. The second issue is a security weakness: although the produced password is 24 characters long, the password is still weak. The weakness comes from the fact that the characters are not being drawn randomly but are based on proximity of keys on a keyboard. The

passwords are not resilient to brute force attack, unless the repetition of the sequences is omitted and the length of the nonrepeating sequence is at least 20 characters. Such a requirement in the length raises a question of whether the purpose of the method—a sequence that is easy to remember—is still met.

We proposed an improvement to the PsychoPass method. First, the user needs to consider the use of the SHIFT and ALT-GR keys in combination with other keys. Second, the keys used need to be 1 or 2 distances apart (not only 1), and third, the number of keys in the sequence shall be at least 9, preferably 10. With the sequence length of 10 characters, there is no need to repeat the sequence as the repetition does not significantly improve the security of the total password. The improved PsychoPass method yields passwords that can only be broken in hundreds of years, considering the current computing powers. The proposed version requires 10 keys, as opposed to 20 keys in the original method, for comparable password strength.

Conflicts of Interest

None declared.

References

1. El Emam K, Moreau K, Jonker E. How strong are passwords used to protect personal health information in clinical trials? *J Med Internet Res* 2011;13(1):e18 [FREE Full text] [doi: [10.2196/jmir.1335](https://doi.org/10.2196/jmir.1335)] [Medline: [21317106](https://pubmed.ncbi.nlm.nih.gov/21317106/)]
2. Cipresso P, Gaggioli A, Serino S, Cipresso S, Riva G. How to create memorable and strong passwords. *J Med Internet Res* 2012;14(1):e10 [FREE Full text] [doi: [10.2196/jmir.1906](https://doi.org/10.2196/jmir.1906)] [Medline: [22233980](https://pubmed.ncbi.nlm.nih.gov/22233980/)]
3. Stallings W. *Cryptography and Network Security, Principles and Practices*, 4th ed. London, UK: Pearson Education; 2006.
4. Hunt T. Our password hashing has no clothes. 2012. URL: <http://www.troyhunt.com/2012/06/our-password-hashing-has-no-clothes.html> [accessed 2012-09-20] [WebCite Cache ID 6AobEryR5]
5. InsidePro. Password Recovery Software. 2012. URL: <http://www.insidepro.com/> [accessed 2012-09-20] [WebCite Cache ID 6AoaUZsx3]
6. Kingsley-Hughes A. Cheap GPUs are rendering strong passwords useless. 2011. URL: <http://www.zdnet.com/blog/hardware/cheap-gpus-are-rendering-strong-passwords-useless/13125> [accessed 2012-09-20] [WebCite Cache ID 6Aob1QCot]
7. Zonenberg A. Distributed hash cracker: A cross-platform gpu-accelerated password recovery system. Troy, NY: Rensselaer Polytechnic Institute; 2009. URL: <http://www.cs.rpi.edu/~zonena/papers/cracker.pdf> [accessed 2013-07-16] [WebCite Cache ID 6IA9PuB7h]
8. Moore GE. Cramming more components onto integrated circuits. *Electronics* 1965;38(8):114-117.
9. Rogers O, Cliff D. A financial brokerage model for cloud computing. *Journal of Cloud Computing* 2012;1(1):1-12.
10. Ladd TD, Jelezko F, Laflamme R, Nakamura Y, Monroe C, O'Brien JL. Quantum computers. *Nature* 2010 Mar 4;464(7285):45-53. [doi: [10.1038/nature08812](https://doi.org/10.1038/nature08812)] [Medline: [20203602](https://pubmed.ncbi.nlm.nih.gov/20203602/)]
11. Lewis MP. *Ethnologue: Languages of the world*. Dallas, TX: SIL International; 2009.
12. Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, Google Books Team, et al. Quantitative analysis of culture using millions of digitized books. *Science* 2011 Jan 14;331(6014):176-182 [FREE Full text] [doi: [10.1126/science.1199644](https://doi.org/10.1126/science.1199644)] [Medline: [21163965](https://pubmed.ncbi.nlm.nih.gov/21163965/)]
13. Project MA. The Biggest MD5 database of Internet. URL: <http://md5.my-addr.com/> [accessed 2012-09-20] [WebCite Cache ID 6AokLW2G5]
14. Blandy. MD5Decrypter. 2012. URL: <http://www.md5decrypter.co.uk/> [accessed 2012-09-20] [WebCite Cache ID 6AokhBgpl]

Abbreviations

AZERTY: specific keyboard layout for the characters of the Latin alphabet that takes its name from the first 6 letters to appear on the first row of alphabetical keys. It is mostly used in France and Belgium.

GPU: graphics processing unit

ICS: Ice Cream Sandwich (Android Version 4.0)

QWERTY: keyboard layout that is mostly used in United Kingdom, United States of America, Canada (and other countries) and is named after the first 6 letters in the first letter row on some keyboards.

QWERTZ: widely used keyboard layout in Central Europe. The name comes from the first 6 letters at the top left of the keyboard.

Edited by G Eysenbach; submitted 24.09.12; peer-reviewed by P Cipresso, R Kok; comments to author 31.01.13; revised version received 04.02.13; accepted 11.06.13; published 13.08.13.

Please cite as:

*Brumen B, Heričko M, Rozman I, Hölbl M
Security Analysis and Improvements to the PsychoPass Method
J Med Internet Res 2013;15(8):e161
URL: <http://www.jmir.org/2013/8/e161/>
doi: [10.2196/jmir.2366](https://doi.org/10.2196/jmir.2366)
PMID: [23942458](https://pubmed.ncbi.nlm.nih.gov/23942458/)*

©Bostjan Brumen, Marjan Heričko, Ivan Rozman, Marko Hölbl. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.08.2013. 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

Analysis of the Security and Privacy Requirements of Cloud-Based Electronic Health Records Systems

Joel JPC Rodrigues^{1*}, BSc, MSc, PhD; Isabel de la Torre^{2*}, BSc, MSc, PhD; Gonzalo Fernández², BSc; Miguel López-Coronado², BSc, MSc, PhD

¹Instituto de Telecomunicações, University of Beira Interior, Covilha, Portugal

²University of Valladolid, Valladolid, Spain

* these authors contributed equally

Corresponding Author:

Joel JPC Rodrigues, BSc, MSc, PhD
Instituto de Telecomunicações, University of Beira Interior
Rua Marques D'Avila e Bolama
Covilha, 6201-001
Portugal
Phone: 351 275242081
Fax: 351 275319899
Email: joeljr@ieee.org

Abstract

Background: The Cloud Computing paradigm offers eHealth systems the opportunity to enhance the features and functionality that they offer. However, moving patients' medical information to the Cloud implies several risks in terms of the security and privacy of sensitive health records. In this paper, the risks of hosting Electronic Health Records (EHRs) on the servers of third-party Cloud service providers are reviewed. To protect the confidentiality of patient information and facilitate the process, some suggestions for health care providers are made. Moreover, security issues that Cloud service providers should address in their platforms are considered.

Objective: To show that, before moving patient health records to the Cloud, security and privacy concerns must be considered by both health care providers and Cloud service providers. Security requirements of a generic Cloud service provider are analyzed.

Methods: To study the latest in Cloud-based computing solutions, bibliographic material was obtained mainly from Medline sources. Furthermore, direct contact was made with several Cloud service providers.

Results: Some of the security issues that should be considered by both Cloud service providers and their health care customers are role-based access, network security mechanisms, data encryption, digital signatures, and access monitoring. Furthermore, to guarantee the safety of the information and comply with privacy policies, the Cloud service provider must be compliant with various certifications and third-party requirements, such as SAS70 Type II, PCI DSS Level 1, ISO 27001, and the US Federal Information Security Management Act (FISMA).

Conclusions: Storing sensitive information such as EHRs in the Cloud means that precautions must be taken to ensure the safety and confidentiality of the data. A relationship built on trust with the Cloud service provider is essential to ensure a transparent process. Cloud service providers must make certain that all security mechanisms are in place to avoid unauthorized access and data breaches. Patients must be kept informed about how their data are being managed.

(*J Med Internet Res* 2013;15(8):e186) doi:[10.2196/jmir.2494](https://doi.org/10.2196/jmir.2494)

KEYWORDS

cloud-computing; eHealth; electronic health records (EHRs); privacy; security

Introduction

Cloud computing environments provide a great opportunity to provide eHealth services in different scenarios in an effective and simple way. The scalability and mobility that a Cloud-based

environment system can offer provides several advantages [1-9], but there are some barriers that must also be managed [10,11]. In the case of deploying a Cloud-based EHR management system, the main advantage is the ability to share patient records with other clinical centers, and the integration of all the EHRs

of a group of clinical centers in order to help medical staff perform their jobs [12-14]. So, how can health care providers and clinical centers guarantee the security, privacy, and confidentiality of their patients' data? The privacy and security of data migrated to the Cloud represents the main barrier that the Cloud computing paradigm must overcome if a Cloud-based eHealth environment is to be deployed. This mission must be performed by both Cloud service providers and health care providers, since hosting EHRs in the Cloud requires a change of approach and they must take into account and address all these risks [15-17].

Security issues are critical when a health care provider plans to deploy a Cloud-based EHR management system. The health care provider must guarantee the security of patient data by ensuring that the Cloud platform has the needed security mechanisms in place. Transmission and network secure protocols also must be deployed in order to avoid external attacks to the data [18]. Moving patient data to the Cloud means that patient files are hosted in the servers of the Cloud service provider [19]. What does this mean? It is essential that these companies ensure the security of their databases so that the data cannot be accessed or modified by unauthorized users. It is important to be aware that privacy and confidentiality terms are essential when EHRs are migrated to the Cloud because of the sensitivity of patient data. In order to avoid unauthorized access, Cloud service providers must deploy authentication systems that ensure the privacy of patient information.

Governments must require that Cloud service providers fulfill the privacy requirements needed to ensure the privacy of patient data. The deployment of a legal framework will help to accomplish a secure environment [13,14]. Privacy policies have been legislated in several countries in order to regulate and safeguard the privacy of patient records. As an example, the US Health Insurance Portability and Accountability Act (HIPAA) regulates the privacy and security of US patient data [20]. These policies depend on each country. Furthermore, EHRs themselves are ruled by standards, which include security and privacy terms, such as Health Level 7 (HL7) [21,22], to guarantee data security and privacy. By combining these standards with Cloud policies and security mechanisms implemented by providers, a secure "Health Cloud" scenario will be achieved.

This paper addresses the health care providers' security and privacy issues that must be considered when deploying EHR management systems. Taking into account these issues on both sides, the migration process will be more secure and transparent. Some security mechanisms necessary to deploy a proper solution are suggested.

We will first elaborate on the issues and requirements for maintaining the security and privacy of EHRs. After that, we explain the requirements that a Cloud-based EHR management system must guarantee in terms of security. Also, some suggestions are given to health care providers in order to facilitate the process.

Methods

For the analysis and study of Cloud-based EHR systems, we reviewed published papers and research about security and privacy issues, which different Cloud computing providers use for development of their Cloud platforms. The related literature was obtained mainly from Medline sources. Direct contact with some Cloud service providers was made. Many publications that show the feasibility of Cloud computing implementations for eHealth services were reviewed in order to look for the latest information on this emerging technology. Most of them show the advantages that Cloud-based solutions can provide to eHealth systems.

Results

Electronic Health Record Security and Privacy Issues

The deployment of EHR management systems is one of the most important achievements in eHealth in recent years. The implementation of these systems has been growing rapidly. In fact, most developed countries have a high level of penetration of this kind of system.

According to Spanish law 41/2002, an EHR is defined as the documentation, which contains information about the clinical evolution of the patient during his or her health assistance process. In this law, the uses of EHRs are set out, requiring medical personnel to maintain the privacy of patients. The Spanish law treats this kind of information as "specially protected" files. This kind of nomenclature is set in the 15/1999 law with the purpose of guarding the privacy of sensitive patient information. The patient's consent is required to manage and access this data, except in the case of an emergency where the patient's life is at risk.

In the United States, HIPAA regulates and establishes the security and privacy requirements of patient data. This law includes two sections on avoiding the improper use of personal information: the Privacy Rule and the Security Rule. The HIPAA Privacy Rule establishes that the Protected Health Information (PHI) must be made available in order to provide the patient medical treatment, either with a Court order or with the authorization of the patient. This rule adds that the entities that use the health information must notify the patient about the use of their PHI. Furthermore, the Privacy Rule requires that entities accessing the PHI use the least amount of patient data necessary to meet their needs. The HIPAA Security Rule was set in 2003 and complements the Privacy Rule, adding several terms to address the digitalization of the patient health information. It has three kinds of security guarantees: administrative, technical, and physical [23-25].

Thus, as outlined above, health care providers must guarantee and preserve the security and privacy of EHRs, and then implement the required security mechanisms to keep patient information safe in the Cloud. Before explaining the mechanisms that a Cloud service provider must implement, we describe the security and privacy requirements of patient records.

Electronic Health Record Security and Privacy Requirements

Before moving EHRs to the Cloud, the EHR systems themselves must set several guarantees to preserve sensitive patient information. The combination of these security requirements with those of the Cloud systems will guarantee the privacy and security of EHRs hosted in the Cloud. The requirements to secure an EHR are described in Table 1 [22]. The security and privacy issues that a Cloud-based system must address in order to safeguard patient files are analyzed in the next section.

Security and Privacy Issues of Cloud-Based Health Solutions

Deploying Cloud-based health solutions is an important step in the development of eHealth. Cloud-based systems allow the ability to create scalable environments, which are adapted to user needs. This total adaptation is complemented by the savings offered by a pay-per-use system, like Cloud computing. Another great advantage comes from the fact that, when EHRs are hosted in the Cloud, medical personnel or patients have the ability to access the information at any time from wherever they have an Internet connection. Currently, with the global economic crisis, saving money could be one of the most important reasons that would drive a company to move its electronic health system into the Cloud. Therefore, Cloud service providers must take advantage of this fact when selling their prospective clients on the advantages of Cloud-based systems.

In order to guarantee the security of their systems, Cloud service providers must install several security mechanisms to keep the safety, privacy, and security of their clients' data. In the section below, we explain the different mechanisms that a Cloud service provider implements in its systems to maintain the security of files in the context of EHR security.

eHealth Cloud Security Issues

A Cloud-based EHR must maintain the same level of data security as data stored in the servers of the health care provider. Patients and medical personnel should know that their personal information is going to be stored with a third-party provider; the provider must guarantee the same security and privacy that the EHRs had in the local servers. The patient, obviously, is not involved in the process of moving their sensitive information to the Cloud, but information should be communicated to patients by the health care providers about the data migration. These communications are not simple notifications; instead, patients should be informed about all the advantages that a Cloud-based system offers for the management of their medical information. Patients should know that data management responsibility lies with both parties: the Cloud service provider and, in a more active way, the health care provider or clinical center. However, there are security issues that should be considered by both providers and customers of a Cloud-based EHR system.

Table 1. Requirements for maintaining the security and privacy of an electronic health record.

Requirements	Description
Authorized access	In order to deploy an authorized-control system, it is essential to deploy an identification system for both patients and health care providers. This identification must be portable between the different entities that have access to the patients' data. This system might be achieved by the ID identifier of each patient. Regarding the authentication, a centralized system based on a public key is viable. A RBAC (Role-Based Access Control) should be deployed in order to allow authorized personnel access to specific data based on their role.
Confidentiality	To guarantee the confidentiality of the communication process, encryption algorithms are used. However, the confidentiality problem in a distributed system arises because it is not possible for the information transmitter system to verify that confidentiality has not been exposed on the receiving end.
Patient's consent	According to the legislation, patients must allow or deny access to their clinical information, except in emergency situations. This consent could be implicit or explicit. Another fact to consider is the need to get access to the EHR-hosted entity from another external one. This process should have the consent of the patient, but in case of emergency, a security mechanism must be provided to avoid this restriction without the patient's consent.
Relevance	All the medical personnel who take part in the diagnostic and treatment process have access to the EHR. Administrative personnel will be able to access the clinical information if their function is relevant to the medical process. Therefore, only the relevant personnel will get access to the patient information. To guarantee that only this level of personnel has been able to access the data, an access control system must be deployed. Given the difficulty of establishing information relevance, it is preferable to have a default permission access and, if necessary, study possible abuses.
Information ownership	The ownership of the EHR is not clearly established. The medical personnel are responsible for this information. However, the patients themselves have the right to access their clinical information.
Information consistency	In an interoperability outline, a correction notification mechanism must be created in order to show changes to the information. This system must allow access to the previous versions of the EHRs, if necessary.
Audits	An audit register should include all accesses to the information and all the changes that have taken place to the EHRs. This system allows the monitoring of access and is a powerful tool to guarantee a secure system. This audit system should fulfill the interoperability requirements.
Archiving	Medical records should be archived for a set period of time, according to the legislation of the respective country. After this period of time, the medical data may be deleted. However, this is not recommended when it comes to EHR management and practice, where the aim is to keep the complete medical information about the patient for his or her lifetime. However, from a logistical standpoint, this would have massive long-term storage requirements.

Role-Based Access

There are many different kinds of personnel who will have access to the patient health record, from the patients themselves to the technicians responsible for the management of the provider’s servers. Physicians, medical personnel, or employees of the Cloud service provider could have access to these data. To ensure the privacy of the patient data, a role-based access system is needed because a doctor may have different access requirements to the patient information than other technical personnel. In order to overcome this problem, an ID code or number must be assigned to each person allowed to access the stored information. Depending on the ID number, the user will belong to a group and each kind of group will have access to a certain part of the patient information [22-26]. For example, patients and doctors will get access to the entire health record whereas the personnel responsible for maintenance of the platform will be able to access only the information they need for proper system operation. With this role-based system, the patients’ privacy is relatively guaranteed. Figure 1 illustrates the different roles that could take part in a Health Cloud and the different versions they will have access to.

Network Security Mechanisms

The main risk to the information will likely be “outside” the Cloud platform. The provider personnel are not the main threat that has to be feared. It is important to know that when moving patient data to the Cloud, health care providers are exposing this information to several external threats because the data are now available via the Internet [23]. Therefore, the responsibility

must lie with the Cloud provider itself to protect the security and privacy of the information by providing the security needed to avoid external attacks to steal or even delete the information.

Data Encryption

All sensitive patient information must be stored securely in a private medical record so that medical information can be shared by different doctors or medical personnel. In order to secure this transaction, the information must be properly encrypted and controlled.

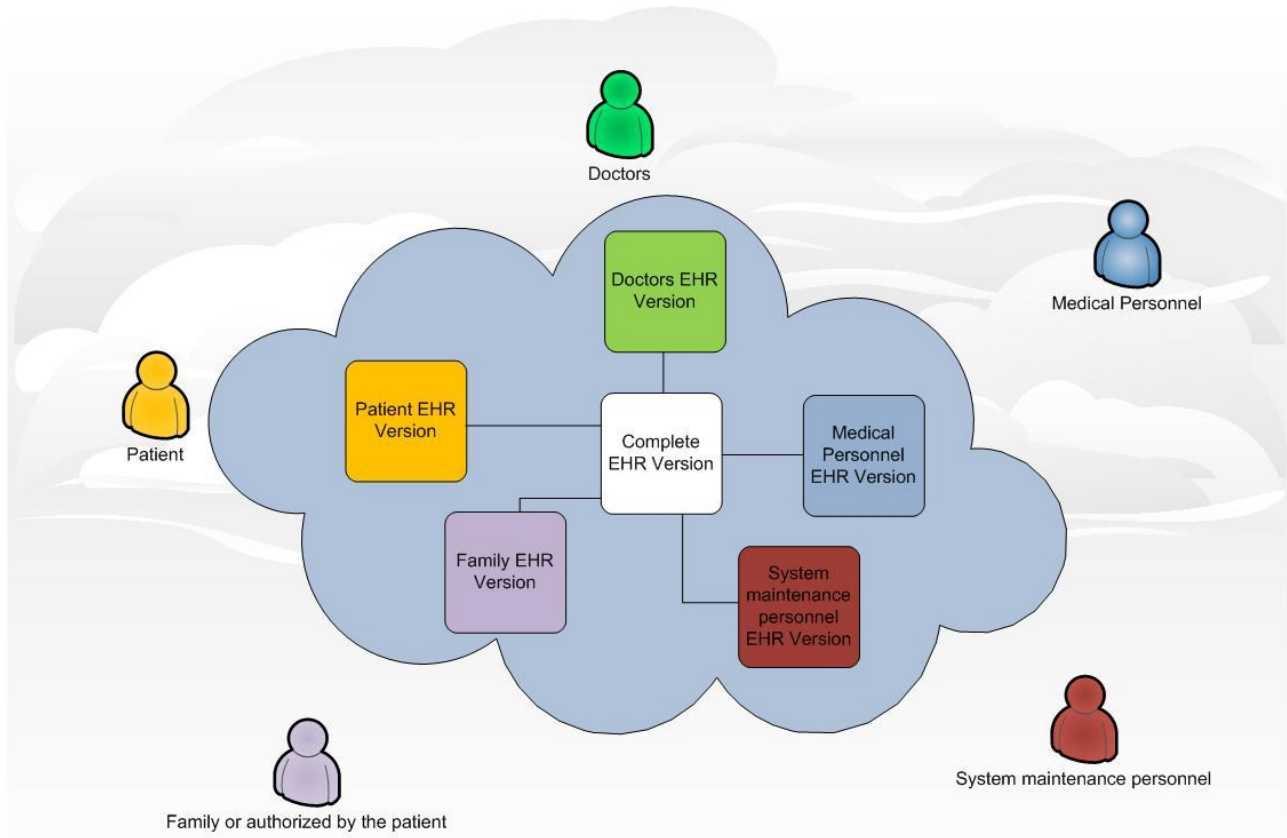
Digital Signature

The digital signature is a very useful tool that provides authenticity, integrity, and nonrepudiation [14-15]. With this security mechanism, the authenticity of the digital record is guaranteed; it will be valuable to deploy this kind of system in the Health Cloud in order to avoid false data transactions. For messages sent through an unsecure channel, the digital signature gives the receiver the reassurance that a message or file was sent by the claimed sender. There are many cryptographic logarithms to deploy this kind of security tool [23].

Monitoring of System Access

Every access to the platform should be monitored in order to create a log of all the people that have had access to the system. In case of an incident, the log can be consulted to solve or find out the cause of the problem. It would be valuable to create a log to track every update and change to each medical record [23].

Figure 1. Role-based system with different electronic health record versions available depending on the kind of user of the Health Cloud.



Suggestions Before Moving Electronic Health Records to the Cloud

The main worries of health care providers planning to move patient information to the Cloud are data security and privacy. Migrating data to the Cloud means that a third party now has control over the Cloud-hosted data. In order to address the risks that could arise, Cloud clients should be well informed before moving data to the Cloud. In order to facilitate this process, the Cloud service provider's customers themselves should be informed about the services the Cloud provider offers them and the security mechanisms installed on the provider's servers. Cloud clients should demand total transparency from the Cloud service provider. Knowing this kind of information is critical to being able to choose the most suitable provider for the client's needs. [Table 2](#) shows several security issues a client should consider when choosing the most appropriate provider [21].

Moving Electronic Health Records to the Cloud: Example of a Cloud Company's Security Requirements

Health care providers that decide to move their EHRs to the Cloud should be aware of these kinds of security mechanisms before migrating their records. There are several well-known Cloud service provider companies, for example, Amazon Web Services, Microsoft Cloud, GoGrid, or Salesforce, with similar security terms as explained below. Thus, this section is useful in the case of choosing a Cloud service provider. Based on the security deployed on several Cloud platforms, we suggest the following mechanisms to secure the Cloud system [22,26,27].

Third-Party Certification

In order to guarantee the safety of the data and meet the requirements of privacy policies, the Cloud provider must be compliant with various certifications and third-party requirements (see [Table 3](#)).

Monitoring

The provider should include automated monitoring tools to provide a high level of service performance and system availability. These tools should be available online for internal and external use.

Notification alarms can be configured when any modification of the data is made by the maintenance personnel or the users themselves. These tools will help track all the information changes made to the stored cloud data. Any kind of incident with the stored data will be monitored.

Information and Communication

In order to use the Cloud platform as a communication channel where personnel could be notified and kept up to date on everything that happens, the Cloud provider should employ various methods of internal communications in order to help employees to understand their roles and responsibilities, and to communicate significant events, if necessary. These communication methods could include orientation and training programs for newly hired personnel, video conferencing, and email, among others.

Employee Lifecycle

Several policies are established in the Cloud platform to manage user access. The Cloud service provider should require that staff with potential access to the patient data undergo an extensive background check (as permitted by law) commensurate with their position and level of data access. Some of these policies are shown in [Table 4](#).

Physical Security

The data center building should be strictly controlled and secured with video surveillance, expert security staff, intrusion detection systems, and other electronic means. The authorized personnel should pass through authentication controls to access the data center floors.

Environmental Safeguards

Innovative architectural and engineering approaches should be used in database centers so as to avoid external agents that could damage them (see [Table 5](#)).

Configuration Management

The company should communicate all updates on both the infrastructure and the software itself, so as to minimize any impact on the customer and the service. The software updating process should be designed to avoid unintended service disruptions and maintain the integrity of service to the customer. Before updating software, these updates should be reviewed, experimented, and approved. The Cloud provider staff would manage the data center infrastructure and be responsible for the hosting management, system scalability, availability and auditing, and security management.

Business Continuity Management

The Cloud service provider must guarantee the availability of the service offered. In order to ensure system availability and continuity, the company should address the security issues considered in [Table 6](#).

Backups

In order to guarantee the existence of the patient data stored in the Cloud, the provider should redundantly store these data. Multiple backups of these data should be stored in different data centers in various locations.

Storage Service Decommissioning

When a Cloud storage service comes to the end of its useful life, the provider should guarantee that data previously stored there is completely removed from its servers. Furthermore, the provider must ensure that unauthorized personnel have not copied these data.

Network Security

The platform itself is not the only element that should be secured by the provider. The Cloud provider must also secure the network. The network provider should guarantee significant protection against traditional network security issues, such as those summarized in [Table 7](#).

Table 2. Suggestions before moving electronic health records to the Cloud.

Security issues	Description
Data security	Because a Cloud provider will have access to all the information concerning the patients, project plans, etc, it is essential to check the provider's reputation in the market. The provider must guarantee that its clients' information would not be misused by any unauthorized personnel. The health care provider should check for the data protection and operational integrity services offered by the provider. Moreover, it is valuable to know the geographic location of the servers where the client data would be hosted. In brief, clients should demand total transparency.
Regulatory compliance	It is important to choose providers with security certifications and are ready for external audits. It is crucial that the provider guarantee the continuity of the service in case the provider has some kind of problem. The client must ensure that the provider operates in the country where the service will be offered. Data logging and data monitoring are important tools that Cloud providers should offer in order to improve the security of the service.
User authentication	Because the data are processed externally by a third party, there is always some inherent risk. The client must know about the personnel who will manage the medical information and what standards for access will be followed by the provider. The client must be informed about the role-based access systems as well as the password handling system configured by the provider.
Data separation	The provider not only handles the data stored in the Cloud but manages the data of other companies who have hired its services. So it is important to know the mechanisms the Cloud provider implements to separate the data of all the companies that are sharing the same servers. The clients must be informed about the availability of the data that the provider guarantees.
Legal issues	A legal framework must guide the policies of the Cloud provider. Intellectual property rights agreements between the two parties should be of prime importance. While the provider owns the right to its infrastructure and applications, the client owns the right to his/her data and computational results.

Table 3. Third-party certifications of the Cloud provider.

Certification	Brief overview
SAS70 Type II	Statement on Auditing Standards No 70: Auditing statement that provides guidance to service auditors when assessing the internal control of a service organization and issuing a service auditor's report.
PCI DSS Level 1	The Cloud provider should be certified with the PCI Data Security Standard as a shared hosting service provider.
ISO 27001	Certification of the Information Security Management System (ISMS) that covers infrastructure, data centers, and service terms.
FISMA	Certification to operate at Federal Information Security Management Act (FISMA) Low Level, which is a US federal law enacted in 2002. It recognizes the importance of information security to the economy and national security interests of the United States.

Table 4. Employee lifecycle policies of a Cloud provider platform.

Policy	Brief overview
Account provisioning	The Cloud provider itself assumes the responsibility of provisioning employees and contractor access. This access to the resources hosted in the Cloud platform must be explicitly approved by the owner or data manager.
Account review	Every access account is reviewed in Cloud platforms every 90 days.
Access removal	Every employee's access account is automatically revoked when it is concluded.
Password policy	Access to the platform is performed by user IDs and passwords to authenticate users to services, resources, and devices, as well as to authorize the appropriate level of access to each user.

Table 5. Environmental safeguards installed in data centers.

Safeguard	Brief overview
Fire detection and suppression	Automatic fire detection and suppression systems are installed in the data center rooms to remove the risk of fire.
Power	24/7 electrical power systems that guarantee the uninterrupted running of the service.
Climate and temperature	In order to prevent overheating of the servers, climate control is required. This is a critical concern for the data center management and consumes lots of energy.
Management	Monitoring systems to control the state of the database equipment.

Table 6. Business continuity management.

Term	Brief overview
Availability	Data centers are built in clusters per regions. In case of failure of one of these data centers, automated processes move the client data traffic away from the affected area.
Incident response	Technical support and coverage to solve any kind of problem 24/7/365 (24 hours a day, 7 days a week, and 365 days a year) must be offered.
Company-wide executive review	A Cloud company should be periodically audited and supported by an internal audit group.

Table 7. Protection against network security issues.

Security network issue	Overview
DDoS attacks	Distributed Denial of Service (DDoS) mitigating techniques is included in the Amazon Web Services (AWS) platform to avoid this kind of attack.
MITM attacks	Man In The Middle (MITM) attacks are avoided because all the endpoints of AWS are secured by Secure Socket Layer (SSL), which provides server authentication.
IP spoofing	Traffic platform is controlled by a firewall infrastructure. Then the stored data cannot send spoofed network data.
Port scanning	Unauthorized port scans by customers are a violation of the provider's use policy. Every reported violation should be investigated.

Discussion

Principal Findings

Migrating electronic health records (EHRs) to the Cloud may represent a great step in the digitalization of medical data. Advantages like scalability, economic model of pay per use, and involving the patient as an active part of the health information management process may assume a change of model in the management of medical records. Several requirements must be taken into account when the time comes to migrate sensitive and private data to the Cloud. Of those requirements, security and privacy of data are the most important ones. In storing the sensitive data of patient health records, Cloud service providers and health care providers must ensure the privacy and confidentiality of the Cloud-hosted data. In order to make this process easier, health care providers, either private or public clinical centers, that have decided to deploy this kind of system, must inform their patients of the change in how their data will be managed and stored. Additionally, a relationship of trust between the health care provider and the Cloud service provider

is an essential factor in this process. In order to achieve this trust, the Cloud provider must guarantee that the security mechanisms are in place to protect the security and privacy of the stored data. An external company is needed to audit the Cloud platform provider in order to show transparency in the management information process. Legislative mechanisms regarding the security of data may be important. Comparing the security terms of several cloud computing companies will be valuable in order to choose the most suitable provider.

Conclusion

With the emergence of Cloud computing, EHR management systems are facing an important platform shift, but such important changes must be approached carefully. In order to make a secure and smooth transition, studying all the security requirements regarding the privacy and confidentiality of patient data are essential. The Cloud computing paradigm is still under development but stands to become revolutionary in many different fields. In the near future, more services and apps will be available, and development will be enhanced.

Acknowledgments

This research has been partially supported by the Spanish Social Security Administration Body (IMSERSO) under the project 85/2010, by the Instituto de Telecomunicações, Next Generation Networks and Applications Group (NetGNA), Portugal, by National Funding from the FCT—Fundação para a Ciência e a Tecnologia through the Pest-OE/EEI/LA0008/2013 Project, and by the AAL4ALL (Ambient Assisted Living for All), project co-funded by COMPETE under FEDER via QREN Programme.

Conflicts of Interest

None declared.

References

1. Furth B, Escalante A. Handbook of Cloud Computing 1st Edition. London: Springer; 2010.
2. Chen YY, Lu JC, Jan JK. A secure EHR system based on hybrid clouds. *J Med Syst* 2012 Oct;36(5):3375-3384. [doi: [10.1007/s10916-012-9830-6](https://doi.org/10.1007/s10916-012-9830-6)] [Medline: [22351166](https://pubmed.ncbi.nlm.nih.gov/22351166/)]

3. Low C, Hsueh Chen Y. Criteria for the evaluation of a cloud-based hospital information system outsourcing provider. *J Med Syst* 2012 Dec;36(6):3543-3553. [doi: [10.1007/s10916-012-9829-z](https://doi.org/10.1007/s10916-012-9829-z)] [Medline: [22366976](https://pubmed.ncbi.nlm.nih.gov/22366976/)]
4. Poulymenopoulou M, Malamateniou F, Vassilacopoulos G. Emergency healthcare process automation using mobile computing and cloud services. *J Med Syst* 2012 Oct;36(5):3233-3241. [doi: [10.1007/s10916-011-9814-y](https://doi.org/10.1007/s10916-011-9814-y)] [Medline: [22205383](https://pubmed.ncbi.nlm.nih.gov/22205383/)]
5. Buyya R, Ranjan R. Special section: Federated resource management in grid and cloud computing systems. *Future Generation Comput Syst* 2010;26(8):1189-1191.
6. Bateman A, Wood M. Cloud computing. *Bioinformatics* 2009;25(12):1475.
7. Armbrust M, Fox A, Griffith R, Joseph AD, Katz R, Konwinski A. A view of cloud computing. *Commun ACM* 2010;53(4):50-58.
8. Creeger M. CTO roundtable: cloud computing. *Commun ACM* 2009;52(8):50-56.
9. Kuo AM. Opportunities and challenges of cloud computing to improve health care services. *J Med Internet Res* 2011;13(3):e67 [FREE Full text] [doi: [10.2196/jmir.1867](https://doi.org/10.2196/jmir.1867)] [Medline: [21937354](https://pubmed.ncbi.nlm.nih.gov/21937354/)]
10. Anderson NR, Lee ES, Brockenbrough JS, Minie ME, Fuller S, Brinkley J, et al. Issues in biomedical research data management and analysis: needs and barriers. *J Am Med Inform Assoc* 2007;14(4):478-488 [FREE Full text] [doi: [10.1197/jamia.M2114](https://doi.org/10.1197/jamia.M2114)] [Medline: [17460139](https://pubmed.ncbi.nlm.nih.gov/17460139/)]
11. Svantesson D, Clarke R. Privacy and consumer risks in cloud computing. *Comput Law Secur Rev* 2010;26(4):391-397.
12. Fernández-Cardeñosa G, De la Torre-Díez I, López-Coronado M, Rodrigues JPC. Analysis of cloud-based solutions on EHRs systems in different scenarios. *J Med Syst* 2012;36(6):3777-3782.
13. Fernández-Cardeñosa G, De la Torre-Díez I, Rodrigues JPC. Analysis of the Cloud Computing Paradigm on Mobile Health Records Systems. In: *Proceedings of the Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. 2012 Presented at: Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing; July 2012; Palermo, Italy.
14. De la Torre-Díez I, Díaz-Pernas FJ, Fernández-Cardeñosa G, Antón-Rodríguez M, Martínez-Zarzuela M, González-Ortega D, et al. Analysis of the benefits and constraints for the implementation of Cloud Computing over an EHRs system. In: *Proceedings of the 6th Euro American Conference on Telematics and Information*. 2012 Presented at: 6th Euro American Conference on Telematics and Information Systems; May 2012; Valencia, Spain.
15. Yellowlees PM, Marks SL, Hogarth M, Turner S. Standards-based, open-source electronic health record systems: a desirable future for the U.S. health industry. *Telemed J E Health* 2008 Apr;14(3):284-288. [doi: [10.1089/tmj.2007.0052](https://doi.org/10.1089/tmj.2007.0052)] [Medline: [18570554](https://pubmed.ncbi.nlm.nih.gov/18570554/)]
16. Blanchet KD. Electronic health records: are consumers riding or driving the car? *Telemed J E Health* 2008 Apr;14(3):210-214. [doi: [10.1089/tmj.2008.9982](https://doi.org/10.1089/tmj.2008.9982)] [Medline: [18570543](https://pubmed.ncbi.nlm.nih.gov/18570543/)]
17. Hargreaves JS. Will electronic personal health records benefit providers and patients in rural America? *Telemed J E Health* 2010 Mar;16(2):167-176. [doi: [10.1089/tmj.2009.0063](https://doi.org/10.1089/tmj.2009.0063)] [Medline: [20082592](https://pubmed.ncbi.nlm.nih.gov/20082592/)]
18. Coles-Kemp L, Reddington J, Williams PAH. Looking at clouds from both sides: The advantages and disadvantages of placing personal narratives in the cloud. *Information Security Technical Report* 2011;16:115-122.
19. Piette JD, Mendoza-Avelares MO, Ganser M, Mohamed M, Marinec N, Krishnan S. A preliminary study of a cloud-computing model for chronic illness self-care support in an underdeveloped country. *Am J Prev Med* 2011 Jun;40(6):629-632. [doi: [10.1016/j.amepre.2011.02.014](https://doi.org/10.1016/j.amepre.2011.02.014)] [Medline: [21565655](https://pubmed.ncbi.nlm.nih.gov/21565655/)]
20. HIPAA General Information. URL: <http://www.cms.gov/Regulations-and-Guidance/HIPAA-Administrative-Simplification/HIPAAGenInfo/index.html?redirect=/HIPAAGenInfo/> [accessed 2013-07-10] [WebCite Cache ID 6I11USrUj]
21. Tejero A, de la Torre I. Advances and current state of the security and privacy in electronic health records: survey from a social perspective. *J Med Syst* 2012 Oct;36(5):3019-3027. [doi: [10.1007/s10916-011-9779-x](https://doi.org/10.1007/s10916-011-9779-x)] [Medline: [21947856](https://pubmed.ncbi.nlm.nih.gov/21947856/)]
22. Saquero-Rodríguez A, De la Torre-Díez I, Durango-Pascual A. Análisis de Aspectos de Interés sobre Privacidad y Seguridad en la Historia Clínica Electrónica. *Revistaesalud.com* 2011;7(27).
23. Zhang R, Liu L. Security Model and Requirements for Healthcare Application Clouds. In: *Proceedings of the IEEE 3rd International Conference on Cloud Computing*. 2010 Presented at: IEEE 3rd International Conference on Cloud Computing; July 2010; Miami, Florida.
24. Cloud Computing: Top 5 Security Concerns. 2012 Feb 23. Health Information Technology, Implementation, Insight, News, Spotlight, Today URL: <http://www.ehrscope.com/blog/cloud-computing-top-5-security-concerns/> [accessed 2013-07-10] [WebCite Cache ID 6I11rOzss]
25. Amazon Web Services: Overview of Security Processes. 2011 May. URL: http://d36cz9buwru1tt.cloudfront.net/pdf/AWS_Security_Whitepaper.pdf [accessed 2013-07-10] [WebCite Cache ID 6I11tBPdk]
26. Force.com Security Resources. URL: <http://wiki.developerforce.com/page/Security> [accessed 2013-07-10] [WebCite Cache ID 6I11uoLos]
27. Fledel Y. Google Android: A Comprehensive Security Assessment. *IEEE Security & Privacy* 2013 (forthcoming). [doi: [10.1109/MSP.2010.2.](https://doi.org/10.1109/MSP.2010.2.)]

Abbreviations

AWS: Amazon Web Services
DDoS: Distributed Denial of Service
EHR: Electronic Health Record
FISMA: Federal Information Security Management Act
HL7: Health Level Seven
HIPAA: Health Insurance Portability and Accountability Act
ISMS: Information Security Management System
MITM: Man In The Middle
PHI: protected health information
SSL: Secure Socket Layer

Edited by K El Emam, G Eysenbach; submitted 15.12.12; peer-reviewed by J Lloret, P Lorenz, C Monteiro; accepted 11.06.13; published 21.08.13.

Please cite as:

*JPC Rodrigues J, de la Torre I, Fernández G, López-Coronado M
Analysis of the Security and Privacy Requirements of Cloud-Based Electronic Health Records Systems
J Med Internet Res 2013;15(8):e186
URL: <http://www.jmir.org/2013/8/e186/>
doi: [10.2196/jmir.2494](https://doi.org/10.2196/jmir.2494)
PMID: [23965254](https://pubmed.ncbi.nlm.nih.gov/23965254/)*

©Joel JPC Rodrigues, Isabel de la Torre, Gonzalo Fernández, Miguel López-Coronado. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 21.08.2013. 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

Attrition Revisited: Adherence and Retention in a Web-Based Alcohol Trial

Elizabeth Murray¹, PhD, FRCGP, FRCP (Edin); Ian R White², PhD; Mira Varaganam³, PhD; Christine Godfrey⁴, BA; Zarnie Khadjesari¹, PhD; Jim McCambridge⁵, PhD

¹e-Health Unit, Research Department of Primary Care and Population Health, University College London, London, United Kingdom

²Medical Research Council Biostatistics Unit, Cambridge, United Kingdom

³Research Department of Primary Care and Population Health, University College London, London, United Kingdom

⁴Department of Health Sciences, University of York, York, United Kingdom

⁵Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, United Kingdom

Corresponding Author:

Elizabeth Murray, PhD, FRCGP, FRCP (Edin)

e-Health Unit

Research Department of Primary Care and Population Health

University College London

Upper Floor 3, Royal Free Hospital

Rowland Hill Street

London, NW3 2PF

United Kingdom

Phone: 44 2077940500 ext 36747

Fax: 44 2077941224

Email: elizabeth.murray@ucl.ac.uk

Related Article:

Letter: <http://www.jmir.org/2014/5/e133/>

Abstract

Background: Attrition is a noted feature of eHealth interventions and trials. In 2005, Eysenbach published a landmark paper calling for a “science of attrition,” suggesting that the 2 forms of attrition—nonusage attrition (low adherence to the intervention) and dropout attrition (poor retention to follow-up)—may be related and that this potential relationship deserved further study.

Objective: The aim of this paper was to use data from an online alcohol trial to explore Eysenbach’s hypothesis, and to answer 3 research questions: (1) Are adherence and retention related? If so, how, and under which circumstances? (2) Do adherence and retention have similar predictors? Can these predictors adequately explain any relationship between adherence and retention or are there additional, unmeasured predictors impacting on the relationship? (3) If there are additional unmeasured predictors impacting on the relationship, are there data to support Eysenbach’s hypothesis that these are related to overall levels of interest?

Methods: Secondary analysis of data from an online trial of an online intervention to reduce alcohol consumption among heavy drinkers. The 2 outcomes were adherence to the intervention measured by number of log-ins, and retention to the trial measured by provision of follow-up data at 3 months (the primary outcome point). Dependent variables were demographic and alcohol-related data collected at baseline. Predictors of adherence and retention were modeled using logistic regression models.

Results: Data were available on 7932 participants. Adherence and retention were related in a complex fashion. Participants in the intervention group were more likely than those in the control group to log in more than once (42% vs 28%, $P<.001$) and less likely than those in the control group to respond at 3 months (40% vs 49%, $P<.001$). Within each randomized group, participants who logged in more frequently were more likely to respond than those who logged in less frequently. Response rates in the intervention group for those who logged in once, twice, or ≥ 3 times were 34%, 46%, and 51%, respectively ($P<.001$); response rates in the control group for those who logged in once, twice, or ≥ 3 times were 44%, 60%, and 67%, respectively ($P<.001$). Relationships between baseline characteristics and adherence and retention were also complex. Where demographic characteristics predicted adherence, they tended also to predict retention. However, characteristics related to alcohol consumption and intention

or confidence in reducing alcohol consumption tended to have opposite effects on adherence and retention, with factors that predicted improved adherence tending to predict reduced retention. The complexity of these relationships suggested the existence of an unmeasured confounder.

Conclusions: In this dataset, adherence and retention were related in a complex fashion. We propose a possible explanatory model for these data.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 31070347; <http://www.controlled-trials.com/ISRCTN31070347> (Archived by WebCite at <http://www.webcitation.org/6IEmNnlCn>).

(*J Med Internet Res* 2013;15(8):e162) doi:[10.2196/jmir.2336](https://doi.org/10.2196/jmir.2336)

KEYWORDS

Internet; eHealth; attrition; adherence; retention; follow-up

Introduction

Background

In a landmark paper published in 2005, Eysenbach [1] argued for a “science of attrition” in the field of eHealth research, noting that attrition is a significant issue in many eHealth studies and calling for researchers to report and explore attrition in eHealth studies. He described 2 forms of attrition: nonusage attrition and dropout attrition. Nonusage attrition, also called low adherence, describes the phenomenon of study participants either not using, or not continuing to use, an eHealth intervention. Under research conditions, nonusage leads to an underestimate of the potential efficacy of the intervention, particularly when the intervention was designed to be used repeatedly over time for maximum effect. Dropout attrition, also called loss to follow-up or low retention, refers to study participants not completing follow-up measures. High loss to follow-up decreases the power of a study, and when extreme, makes it hard to interpret the results of a study because there is no way of knowing what effect the intervention had in those who did not provide follow-up data.

To identify papers responding to Eysenbach’s call, we undertook a search of PubMed for papers published between 2005 and the end of 2011 that addressed either form of attrition. The search strategy combined the concepts of Web-based interventions with attrition (either nonusage attrition/adherence or dropout attrition/retention). Of 2581 unique citations, more than 60 papers reported relevant information, including 5 systematic reviews [2-6], 18 trials determining the effects of a specific intervention on either adherence or retention, 3 qualitative studies exploring participant reasons for adherence, 19 studies reporting secondary analyses of data from trials or cohort studies, and a number of studies that used a range of other methodologies. Most (n=45) focused on factors associated with adherence to the intervention, whereas 11 focused on trial retention and 4 looked at the relationship between adherence and retention.

Adherence to the Intervention (Nonusage Attrition)

Adherence to any specified intervention may be related to characteristics of the intervention, characteristics of the user, or characteristics of the condition addressed by the intervention. Characteristics of the intervention that may improve adherence to the intervention include a strong theoretical foundation [7], perceived personal relevance to the user [8,9], perceived

effectiveness [10,11], tailoring [12,13], persuasive technologies [3], credibility [14,15], social networking [16,17], and regular push factors including human support [18-20] and/or periodic prompts either by email or telephone [6]. There is conflicting evidence on adherence and characteristics of the user. Although many researchers have found that women, older people, and well-educated people are more likely to demonstrate adherence to Web-based interventions than males, younger people, and less-educated people [8,9,11,12], others have found no association between adherence and age, gender, or education [21]. As stated by Melville et al [4] in their review of literature exploring the variables associated with adherence to Internet programs for psychological disorders: “Despite the numerous variables explored, evidence on any specific variables that may make an individual more likely to drop out of Internet-based treatment is currently limited.” We were unable to identify data comparing adherence to similar interventions for different health conditions or health behaviors.

Study Retention (Dropout Attrition)

Improving retention to studies has received less attention in the eHealth literature than improving adherence to interventions. For online questionnaires, the appearance, order, relevance, length, and origin of the questionnaire all seem important [22-24]. Incentives may improve response rates, but may have to have considerable value before having an impact [25]. Providing feedback on questionnaires may improve response rates [26]. There has also been some work characterizing participants who are more likely to drop out from studies, with better response rates reported for people who are white, older, better educated, with good Internet skills [27-31]. Bull et al [32] have had considerable success in improving retention rates in trials of online sexual health promotion. After early difficulties with a trial that only managed 15% retention at follow-up [32], they amended their approach and through a combination of automated electronic and personalized approaches to increase and diversify recruitment, verify participant eligibility and increase retention, achieved a 79% follow-up rate at 1 month [33]. Although eHealth researchers have paid less attention to study retention than to intervention adherence, the general methodological literature has a great deal of information on improving retention, much of which applies to online studies as well as offline ones [34-37].

Relationship Between Adherence and Retention

In Eysenbach's original paper, he posited that the 2 forms of attrition were related to one another by a single underlying mechanism—losing interest—and called for empirical studies to test this hypothesis [1].

In a systematic review and meta-analysis of papers published before 2009, Cugelman et al [3] set out to explore 3 outcomes: (1) the overall effect size of online interventions on voluntary behaviors, (2) the impact of various intervention components designed to influence behavior, and (3) the relationship between dose (exposure) of intervention, effect of intervention, and study retention. Thirty-one papers were included in the review. The authors concluded that, despite 1 contradictory correlation, the evidence suggested that intervention adherence was positively correlated with behavioral change. Only 5 (unspecified) studies could be used to assess the relationship between intervention adherence and study retention, but this analysis did reveal a significant positive correlation between them [3]. Subsequently, Couper et al [13] showed that engagement was significantly associated with completion of follow-up data in a randomized controlled trial of alternative versions of an online intervention to promote consumption of fruit and vegetables.

In summary, although many researchers have provided data contributing to a science of attrition, there remain many unanswered questions:

1. Are adherence and retention related? If so, how, and under which circumstances?
2. Do adherence and retention have similar predictors? Can these predictors adequately explain any relationship between adherence and retention or are there additional, unmeasured predictors influencing the relationship?
3. If there are additional unmeasured predictors affecting the relationship, are there data to support Eysenbach's hypothesis that these are related to overall levels of interest?

This paper addresses these questions through secondary analysis of an Internet-based trial of an online intervention to help heavy drinkers reduce their alcohol consumption.

Methods

Ethical Approval

Ethical approval was obtained from University College London ethics committee.

Study Procedures and Participants

The data reported here were taken from the Down Your Drink (DYD) randomized controlled trial (DYD-RCT; ISRCTN: 31070347), a trial of an online intervention to help hazardous or harmful drinkers reduce their alcohol consumption [38]. Study participants were adults who had browsed the Web and found the DYD home page, which invited them to “find out if you are drinking too much” by completing a brief 3-item screening test, the consumption questions of the Alcohol Use Disorders Identification Test (AUDIT-C) [39]. Those whose score indicated they were drinking at hazardous or harmful levels were invited to take part in the trial (AUDIT-C scores ≥ 5) [40]. Potential participants were informed that the trial

compared 2 websites to see which was more effective in helping users reduce their alcohol consumption. To ensure participants were aware that they were participating in a research study, the informed consent procedures required users to navigate sequentially through 11 pages, provide online consent, create a username and password, and undergo an email validation stage. After their email had been validated, users had to complete baseline data questionnaires, including demographic data, a log of past-week alcohol consumption (the TOT-AL) [41], a 5-item health-related quality of life measure (the EQ-5D) [42], and 1 of 4 randomly allocated secondary outcome measures. Only after all baseline data were completed were participants randomly allocated to either the intervention or the control group. Participants were automatically routed to the website they had been allocated to, so all participants visited their allocated site at least once.

Intervention and Comparator

The intervention website was a theoretically informed website with multiple interactive features. It contained 3 phases. Phase 1 (“It's Up to You”) used the principles of motivational interviewing to enable the user to reach a considered decision on whether to change their alcohol consumption. Phase 2 (“Making the Change”) used cognitive behavioral therapy techniques to help users reduce their consumption, whereas Phase 3 (“Keeping on Track”) focused on relapse prevention, a further cognitive behavioral approach. Interactive e-tools, such as the drinking episode diary, provide opportunities for users to reflect on the role alcohol plays in their life and consider alternatives [43]. Users were free to use the intervention website in any way they wanted, but there was an expectation among the developers that users would log on repeatedly to use all elements of the program. The comparator website had a similar look and feel in terms of colors, graphics, and tone. It presented simple information about the harms of excess alcohol consumption, with untailed advice on how to cut down. It had no interactive tools and no drinking diary.

Follow-Up

The primary outcome point was 3 months after randomization. Data collected at follow-up included past-week alcohol consumption (the primary outcome), the EQ-5D, and the same secondary outcome measure completed at baseline. Data were collected online, with participants sent an email request for follow-up data. The email contained an embedded hotlink that led through to the questionnaires. Up to 3 reminders were sent at 7-day intervals to nonresponders, with the final reminder asking participants to tell us their total past-week alcohol consumption only. This follow-up regime was selected on the basis of our pilot study, exploring different methods of optimizing retention [40].

Measures of Attrition and Retention

For the purposes of this paper, the 2 outcomes of interest were adherence to the intervention and retention in the trial. Adherence to the intervention was categorized by number of log-ons to the site into 3 groups: users who logged in once only, users who logged in twice, and users who logged in 3 or more times. This categorization was empirically based because the

content was highly person-centered so that there were no recommended patterns of use. Retention was treated as a binary variable: participants either did or did not provide follow-up data at 3 months.

The independent variables were the data collected at baseline. Demographic variables were age, gender, educational status (categorized as having vs not having a university degree), marital status (categorized as married/living with partner vs single), having children (categorized as 1 or more vs none), ethnicity (white British or anything else), country of residence (Britain vs rest of the world), and providing offline contact details (provided either address or phone number vs not providing any offline details). Clinical variables were past-week alcohol consumption in units (in which 1 unit is equivalent to 8 g of ethanol), EQ-5D scores, and scores on 2 single-item measures of confidence and intention, both scored from 1 to 5 with 5 indicating the highest level of confidence or intention. For these items, participants were asked, "How confident are you in your ability to reduce your drinking?" and "How strong would you rate your intention to reduce your drinking in the next 3 months?" The EQ-5D scores were obtained from the 5-item questionnaire following the standard procedures [44].

Statistical Methods

Predictors of adherence and retention were modeled by using logistic regression models with outcomes whether a user (1) logged in twice or more, (2) logged in 3 times or more, and (3) responded at 3 months. To explore the association between adherence and retention, dummy variables for exactly 2 log-ins and 3 or more log-ins were included in model 3.

To select a set of independent variables for all adjusted analyses, we first fitted models 1 to 3 using all the demographic and clinical baseline variables listed previously as independent variables. Past-week alcohol consumption was log-transformed after adding 1 unit/week. For each independent variable, we found the smallest *P* value across models 1 to 3 and we dropped the independent variable with the largest value of this smallest *P* value. To focus on stronger predictors, we repeated this procedure until each independent variable was significant at $P < .01$ in at least 1 of the 3 models. An interaction between TOT-AL and gender was included to allow for women's typically lower levels of drinking. Goodness of fit was assessed by the Hosmer-Lemeshow test [45]. To facilitate comparisons of the effects of different independent variables, regression coefficients for quantitative independent variables were expressed per 1 standard deviation change. One individual with missing ethnicity at baseline and 2 with missing TOT-AL at baseline were omitted from this analysis.

Results

Participants

Data were available on 7932 participants. Demographic and alcohol-related characteristics at baseline are presented in Table 1. The mean age was 38 years, more than half were female, and over half had a university degree. Participants were drinking heavily, with a mean past-week alcohol consumption of 57 units/week. Most participants were intending to reduce their

alcohol consumption with a mean score of 3.8 on a scale of 1 to 5; however, they were less confident about their ability to reduce consumption (mean score 2.8).

Adherence and Retention at Three Months

Of the 7932 participants, 5165 (65%) logged in once only, 1538 (19%) logged in twice, and 1229 (16%) logged in 3 or more times (Table 2). Participants in the intervention group were more likely than those in the control group to log in more than once (42% vs 28%, $P < .001$). These adherence rates were much lower than anticipated for the intervention group, with relatively few users making repeated visits.

Retention also varied by arm. The overall response rate at 3 months was 45% (3528/7932). Participants in the intervention arm were less likely to respond than those in the control arm (40% vs 49%, $P < .001$).

These data appeared to conflict with Eysenbach's hypothesis of a common factor in adherence and retention because participants in the intervention arm were both more likely to log in to the intervention and less likely to respond to follow-up than those in the control group. However, within each arm, the data were supportive of the Eysenbach hypothesis, with participants who logged in more frequently being more likely to respond at 3 months (Table 2). The overall difference in response rates at 3 months between those with 2 and ≥ 3 log-ins was smaller than the difference within randomized groups, because this association is confounded by randomized groups. Multivariate analysis confirmed that adherence and retention were positively correlated, with participants who logged in more often being more likely to respond at 3 months (Table 2).

These data suggest that the answer to our first research question (are adherence and retention related?) is yes. In this dataset, adherence and retention are related, but the relationship is not straightforward. Overall, participants who logged in more frequently were more likely to respond at 3 months, but those in the intervention arm were both more likely to log in more than once and less likely to respond than those in the control arm.

Determinants of Attrition and Retention

Our second research question was "Do adherence and retention have similar determinants? Can these determinants adequately explain any relationship between adherence and retention or are there additional unmeasured determinants impacting on the relationship?"

To address this question, we first explored baseline predictors of adherence and retention (Table 3). Two variables were dropped by the variable selection procedure: marital status (which was correlated with children) and country of residence (which was correlated with ethnicity). The 3 fitted models all had adequate goodness of fit (all *P* values $> .05$).

Where demographic characteristics were found to predict adherence, they tended to also predict retention and vice versa, although the relationships did not always reach statistical significance. Thus, being older, female, having a university degree, and not having children were all predictive of being

more likely to log in more frequently and provide follow-up data at 3 months (Table 3).

In contrast, the relationship between alcohol-related characteristics including past-week consumption, intention to reduce consumption, and level of confidence in the ability to reduce consumption, and adherence or retention went in opposite directions (Table 3). Heavier drinkers, those with higher intention to reduce their drinking, and those with lower confidence in their ability to do so were more likely to log in more often and less likely to respond at 3 months. There was an interaction between amount of alcohol consumed and gender, such that for women, heavier drinking at baseline was associated with greater likelihood of logging on 2 or more times but reduced likelihood of response, but this relationship was not seen for men.

Table 3 also shows that the factor with most impact on adherence and retention was allocation to intervention or control.

These data suggest that the answer to our second research question was that although there was some similarity between the determinants of adherence and the predictors of retention, the relationship between adherence and retention could not be wholly explained by the predictors measured at baseline, suggesting there were additional unmeasured confounders affecting this relationship.

Our third research question was whether the data supported Eysenbach’s hypothesis that participant’s overall level of interest was the factor responsible for any relationship between adherence and retention. As discussed previously, the data pertaining to this question were somewhat contradictory, with adherence and retention positively correlated within each arm of the trial but not across arms because participants in the intervention arm were more likely to adhere but less likely to respond than participants in the control arm.

Table 1. Baseline characteristics of the study population (N=7932).

Characteristic	Intervention (n=3970)	Control (n=3962)
Age (years), mean (SD)	37.97 (10.96)	38.29 (10.78)
Gender (female), n (%)	2246 (57)	2299 (58)
Have university degree, n (%)	2067 (52)	2026 (51)
White British, n (%)	3317 (84)	3316 (84)
Have children, n (%)	2052 (52)	2027 (51)
Provided offline address or telephone number, n (%)	1559 (39)	1528 (39)
Past-week alcohol consumption in units ^a , mean (SD)	57.68 (39.62)	56.86 (38.09)
EQ-5D, mean (SD)	0.84 (0.19)	0.84 (0.19)
Confidence ^b , mean (SD)	2.77 (1.16)	2.79 (1.15)
Intention ^b , mean (SD)	3.82 (1.09)	3.85 (1.06)

^a 1 unit = 8 g ethanol.

^b Confidence and intention scored on a 5-point scale with 1 as the lowest and 5 as the highest.

Table 2. Adherence and retention.

Number of log-ins (adherence)	Overall (N=7932)			Intervention (n=3970)			Control (n=3962)		
	n (%)	Responded at 3 months (retention), n (%)	Adjusted OR (95% CI)	n (%)	Responded at 3 months (retention), n (%)	Adjusted OR (95% CI)	n (%)	Responded at 3 months (retention), n (%)	Adjusted OR (95% CI)
1	5165 (65)	2036 (39)	Ref	2324 (59)	792 (34)	Ref	2841 (72)	1244 (44)	Ref
2	1538 (19)	816 (53)	1.79 (1.59, 2.01) ^a	745 (19)	343 (46)	1.79 (1.50, 2.13) ^a	793 (20)	473 (60)	1.88 (1.60, 2.21) ^a
3+	1229 (16)	676 (55)	1.92 (1.68, 2.18) ^a	901 (23)	456 (51)	2.12 (1.80, 2.50) ^a	328 (8)	220 (67)	2.58 (2.02, 3.31) ^a

^aP<.001

Table 3. Baseline predictors of adherence and retention.

Characteristic	Logged in twice OR (95% CI)	<i>P</i>	Logged in ≥ 3 times OR (95% CI)	<i>P</i>	Responded at 3 months OR (95% CI)	<i>P</i>
Age^a						
Per 11-year increase	1.23 (1.16-1.30)	<.001	1.41 (1.31-1.52)	<.001	1.36 (1.29-1.44)	<.001
Gender (at logtotal + 1 = 3.83)						
Females vs males	1.12 (1.02-1.23)	.02	1.12 (0.83-1.28)	.09	1.35 (1.22-1.47)	<.001
University degree						
Degree vs no degree	1.24 (1.13-1.37)	<.001	1.31 (1.15-1.50)	<.001	1.17 (1.07-1.29)	.001
Ethnicity						
White British vs other	0.89 (0.78-1.01)	.08	0.85 (0.72-1.00)	.06	1.24 (1.09-1.41)	.001
Children						
No children vs children	1.12 (1.01-1.25)	.04	1.15 (0.99-1.33)	.06	1.25 (1.12-1.39)	<.001
Provided address or phone number						
Yes vs no	1.1 (0.91-1.11)	.88	0.91 (0.79-1.03)	.14	1.20 (1.09-1.32)	<.001
Past-week alcohol consumption in units (logtotal + 1)						
Per 0.78-unit increase in women ^a	1.13 (1.05-1.21)		1.15 (1.03-1.27)		0.87 (0.81-0.93)	
Per 0.78-unit increase in men ^a	0.99 (0.93-1.07)		0.97 (0.88-1.07)		0.91 (0.85-0.97)	
EQ5D						
Per 0.19-unit increase ^a	1.02 (0.97-1.07)	.40	0.99 (0.93-1.06)	.80	1.07 (1.02-1.12)	.01
Confidence						
Per 1.15-unit increase ^a	0.92 (0.87-0.97)	.001	0.93 (0.87-1.0)	.04	1.06 (1.01-1.12)	.013
Intention						
Per 1.08-unit increase ^a	1.27 (1.21-1.34)	<.001	1.45 (1.35-1.56)	<.001	0.86 (0.82-0.90)	<.001
Trial allocation group						
Intervention vs control	1.84 (1.68-2.03)	<.001	3.43 (2.99-3.94)	<.001	0.70 (0.64-0.76)	<.001

^a Continuous predictors expressed as per 1 SD change.

Discussion

Main Findings

Participants in the intervention arm were more likely to use the intervention and less likely to respond to requests for follow-up data at 3 months than those in the control arm. This relationship is likely to be causal because these data were obtained in a RCT in which the only difference between groups was the allocated intervention. We can conclude, therefore, that allocation to the intervention rather than control led to lower levels of follow-up.

Within each trial arm, there was a strong association between logging in more and being more likely to provide follow-up data at 3 months. As expected, the relationship between demographic variables, adherence, and retention tended to be in the same direction, in which factors associated with greater adherence were also associated with greater retention. However, the variables pertaining to alcohol consumption (past-week alcohol consumption, intention to reduce consumption, and confidence in ability to reduce consumption) tended to impact

on adherence and retention in opposite directions. Participants who may be more likely to benefit from the intervention (heavier drinkers, drinkers with greater intention to reduce, and those with less confidence in their ability to reduce) made greater use of both intervention and control websites (higher adherence) but were less likely to provide follow-up data.

These data provide some answers to 2 of the 3 research questions posed. In this study, adherence and retention were related. Although there was some similarity between the predictors of adherence and the predictors of retention, the relationship between adherence and retention could not be wholly explained by the predictors measured at baseline, suggesting there were additional unmeasured confounders affecting this relationship.

Potential Explanatory Model

Our third research question was whether any additional factors influencing the relationship were related to overall levels of interest as postulated by Eysenbach. Our data cannot directly address this question, but 1 possible explanatory model is presented in [Figure 1](#). This model posits that allocation to the

active or comparator intervention has a causal relationship with both adherence and retention. Allocation to the active intervention increased adherence to the intervention compared to the control and decreased trial retention. We can speculate that this may reflect at least some people in the intervention arm perceiving the website as helpful and, hence, continuing to use it until they felt they had received an adequate “dose” of intervention. Some participants reported being unable to distinguish between “intervention content” and “trial content” [46]. Therefore, when requests for follow-up data were received, some participants in the intervention arm were more likely to feel they had a sufficient dose and more likely to ignore the requests for follow-up data, even though there was no evidence in the trial that they had changed their behavior any more than participants in the control arm had. In contrast, some people allocated to the control may not have perceived it to be useful because it was not designed to be effective and, thus, continued to have unmet need. Requests for follow-up data may have been seen as an opportunity to monitor and reflect on their alcohol consumption—an opportunity which those in the control group welcomed and were more likely to complete the outcome measures.

The data also suggest that there are user characteristics which influence adherence that act in different ways. We hypothesize that demographic factors are indicative of an unmeasured or latent variable, which we could describe as a “propensity to comply.” Participants with higher propensity to comply are more likely to use the intervention more often and to respond to requests for follow-up data. The propensity to comply may overwhelm the effect of being allocated to the intervention arm on retention. In contrast, alcohol-related factors, including past-week alcohol consumption, intention to reduce consumption, and confidence in one’s ability to reduce consumption, increase use of the intervention and reduce retention.

There are, of course, other possible interpretations of the data. One alternative interpretation is that participants allocated to the active intervention were more frustrated by the gap between their expectations of the intervention and their experience of

using it. The active site promised a complete suite of tools to help users make a decision and then act on it, but the low number of log-ons clearly suggested that most users did not use it as planned. It is possible that this disappointment diminished willingness to have contact with the researchers, leading to lower retention. In contrast, the control website only offered straightforward information and, thus, produced less disappointment, leading to greater openness to contact with the researchers, particularly where there were unmet needs as represented by unsuccessful attempts at behavior change.

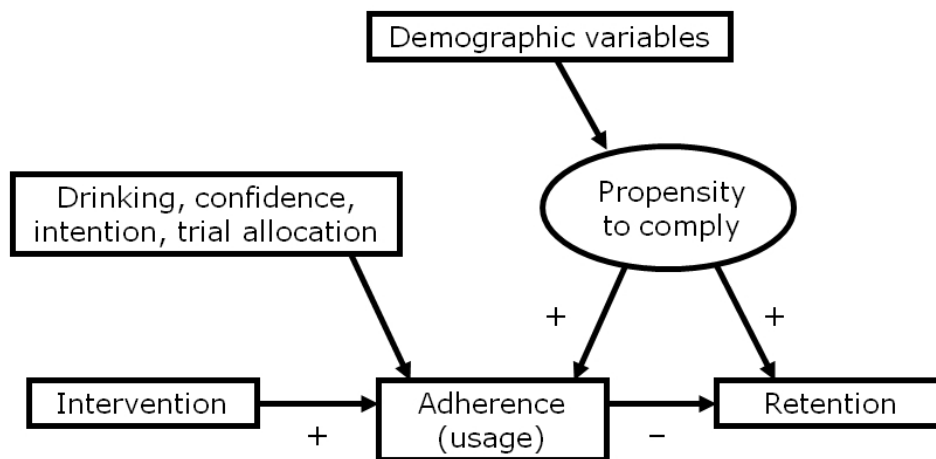
Relationship to Previous Literature

Our data build on the available literature in this field, and may help explain some of the contradictory results seen previously. Previous papers have looked for simple relationships between demographic factors, such as age, gender, or education, and either adherence or retention. Our results suggest that a more complex model is required which takes other factors into account. It is also likely that these relationships will vary according to the population and the behavior or condition studied.

Methodological Strengths and Weaknesses

There are many strengths to these data. They were derived from a large online RCT that had automated randomization ensuring complete concealment of allocation and automated data collection procedures that ensured that all data obtained were of adequate quality for analysis. The large sample size of nearly 8000 participants allowed for highly powered multivariate analyses. The main weaknesses are related to this being a secondary analysis. The initial trial was not designed to address the research questions posed in this paper. We had no a priori definition of adherence and used number of log-ins as the simplest measure of adherence because previous data have shown that number of log-ins and number of pages visited are highly correlated [47]. If there had been a prescribed way of using the intervention we could have looked for adherence to this, but the site was designed to be used differently by different users according to their needs so that participants would make use of the sections or components they found helpful.

Figure 1. Possible model to explain relationship between adherence and retention.



Implications

The data presented here can only be thought of as hypothesis generating. Clearly, further studies specifically designed to test this model and related questions are needed before any firm conclusions can be drawn. However, if confirmed in other studies, these data have implications for both intervention and trial design. Because the biggest predictor of adherence was allocation to the intervention rather than the control arm, researchers need to focus on ensuring that Web-interventions are attractive to the user. The literature suggests that this can best be done by ensuring that the intervention is theoretically informed, has strong perceived personal relevance and effectiveness for users (eg, through tailoring and ensuring credibility), and multiple push factors, both automated (eg, email or short message service text prompts) or human (eg, facilitation or coaching). Previous authors have suggested that it may be possible to identify subgroups of the population who

are most likely to adhere [9]. Our data suggest this is unlikely to be fruitful.

The implications for trial design are more challenging to elaborate. If our hypothesis that low retention was related to users in the intervention group feeling that their needs had been met were to be confirmed, this potentially has profound implications for the design of Web-based trials of Web-based interventions.

Conclusions

In an online RCT of a Web-based intervention to help hazardous and harmful drinkers reduce their alcohol consumption, adherence and retention were related in a complex manner. Some user characteristics, particularly demographic variables, had a positive impact on both adherence and retention, whereas behavioral and related variables increased adherence and reduced retention. We have proposed various possible hypotheses to guide further study.

Conflicts of Interest

None declared.

References

1. 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/)]
2. 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/)]
3. Cugelman B, Thelwall M, Dawes P. Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors. *J Med Internet Res* 2011;13(1):e17 [[FREE Full text](#)] [doi: [10.2196/jmir.1367](https://doi.org/10.2196/jmir.1367)] [Medline: [21320854](https://pubmed.ncbi.nlm.nih.gov/21320854/)]
4. 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/)]
5. 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/)]
6. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. *J Med Internet Res* 2009;11(2):e16 [[FREE Full text](#)] [doi: [10.2196/jmir.1138](https://doi.org/10.2196/jmir.1138)] [Medline: [19632970](https://pubmed.ncbi.nlm.nih.gov/19632970/)]
7. 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/)]
8. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. *J Med Internet Res* 2008;10(5):e36 [[FREE Full text](#)] [doi: [10.2196/jmir.1002](https://doi.org/10.2196/jmir.1002)] [Medline: [18984557](https://pubmed.ncbi.nlm.nih.gov/18984557/)]
9. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, et al. The ins and outs of an online bipolar education program: a study of program attrition. *J Med Internet Res* 2010 Dec;12(5):e57 [[FREE Full text](#)] [doi: [10.2196/jmir.1450](https://doi.org/10.2196/jmir.1450)] [Medline: [21169169](https://pubmed.ncbi.nlm.nih.gov/21169169/)]
10. Gerhards SA, Abma TA, Arntz A, de Graaf LE, Evers SM, Huibers MJ, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. *J Affect Disord* 2011 Mar;129(1-3):117-125. [doi: [10.1016/j.jad.2010.09.012](https://doi.org/10.1016/j.jad.2010.09.012)] [Medline: [20889214](https://pubmed.ncbi.nlm.nih.gov/20889214/)]
11. Zbikowski SM, Hapgood J, Smucker Barnwell S, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. *J Med Internet Res* 2008;10(5):e41 [[FREE Full text](#)] [doi: [10.2196/jmir.999](https://doi.org/10.2196/jmir.999)] [Medline: [19017583](https://pubmed.ncbi.nlm.nih.gov/19017583/)]
12. Wangberg SC, Bergmo TS, Johnsen JA. Adherence in Internet-based interventions. *Patient Prefer Adherence* 2008;2:57-65 [[FREE Full text](#)] [Medline: [19920945](https://pubmed.ncbi.nlm.nih.gov/19920945/)]
13. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and retention: measuring breadth and depth of participant use of an online intervention. *J Med Internet Res* 2010;12(4):e52 [[FREE Full text](#)] [doi: [10.2196/jmir.1430](https://doi.org/10.2196/jmir.1430)] [Medline: [21087922](https://pubmed.ncbi.nlm.nih.gov/21087922/)]

14. Nordgreen T, Havik OE, Ost LG, Furmark T, Carlbring P, Andersson G. Outcome predictors in guided and unguided self-help for social anxiety disorder. *Behav Res Ther* 2012 Jan;50(1):13-21. [doi: [10.1016/j.brat.2011.10.009](https://doi.org/10.1016/j.brat.2011.10.009)] [Medline: [22134140](https://pubmed.ncbi.nlm.nih.gov/22134140/)]
15. Harris PR, Sillence E, Briggs P. The effect of credibility-related design cues on responses to a web-based message about the breast cancer risks from alcohol: randomized controlled trial. *J Med Internet Res* 2009;11(3):e37 [FREE Full text] [doi: [10.2196/jmir.1097](https://doi.org/10.2196/jmir.1097)] [Medline: [19709989](https://pubmed.ncbi.nlm.nih.gov/19709989/)]
16. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. *J Med Internet Res* 2010;12(4):e71 [FREE Full text] [doi: [10.2196/jmir.1338](https://doi.org/10.2196/jmir.1338)] [Medline: [21169160](https://pubmed.ncbi.nlm.nih.gov/21169160/)]
17. Stoddard JL, Augustson EM, Moser RP. Effect of adding a virtual community (bulletin board) to smokefree.gov: randomized controlled trial. *J Med Internet Res* 2008 Dec;10(5):e53 [FREE Full text] [doi: [10.2196/jmir.1124](https://doi.org/10.2196/jmir.1124)] [Medline: [19097974](https://pubmed.ncbi.nlm.nih.gov/19097974/)]
18. Glasgow RE, Kurz D, King D, Dickman JM, Faber AJ, Halterman E, et al. Outcomes of minimal and moderate support versions of an internet-based diabetes self-management support program. *J Gen Intern Med* 2010 Dec;25(12):1315-1322 [FREE Full text] [doi: [10.1007/s11606-010-1480-0](https://doi.org/10.1007/s11606-010-1480-0)] [Medline: [20714820](https://pubmed.ncbi.nlm.nih.gov/20714820/)]
19. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10939 [FREE Full text] [doi: [10.1371/journal.pone.0010939](https://doi.org/10.1371/journal.pone.0010939)] [Medline: [20544030](https://pubmed.ncbi.nlm.nih.gov/20544030/)]
20. Robinson E, Titov N, Andrews G, McIntyre K, Schwencke G, Solley K. Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One* 2010;5(6):e10942 [FREE Full text] [doi: [10.1371/journal.pone.0010942](https://doi.org/10.1371/journal.pone.0010942)] [Medline: [20532167](https://pubmed.ncbi.nlm.nih.gov/20532167/)]
21. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. *J Med Internet Res* 2011;13(1):e9 [FREE Full text] [doi: [10.2196/jmir.1391](https://doi.org/10.2196/jmir.1391)] [Medline: [21371992](https://pubmed.ncbi.nlm.nih.gov/21371992/)]
22. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. *J Med Internet Res* 2011;13(4):e96 [FREE Full text] [doi: [10.2196/jmir.1733](https://doi.org/10.2196/jmir.1733)] [Medline: [22100793](https://pubmed.ncbi.nlm.nih.gov/22100793/)]
23. Ekman A, Klint A, Dickman PW, Adami HO, Litton JE. Optimizing the design of web-based questionnaires--experience from a population-based study among 50,000 women. *Eur J Epidemiol* 2007;22(5):293-300. [doi: [10.1007/s10654-006-9091-0](https://doi.org/10.1007/s10654-006-9091-0)] [Medline: [17206467](https://pubmed.ncbi.nlm.nih.gov/17206467/)]
24. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010;11:2 [FREE Full text] [doi: [10.1186/1745-6215-11-2](https://doi.org/10.1186/1745-6215-11-2)] [Medline: [20064225](https://pubmed.ncbi.nlm.nih.gov/20064225/)]
25. Khadjesari Z, Murray E, Kalaitzaki E, White IR, McCambridge J, Thompson SG, et al. Impact and costs of incentives to reduce attrition in online trials: two randomized controlled trials. *J Med Internet Res* 2011;13(1):e26 [FREE Full text] [doi: [10.2196/jmir.1523](https://doi.org/10.2196/jmir.1523)] [Medline: [21371988](https://pubmed.ncbi.nlm.nih.gov/21371988/)]
26. Bälter O, Fondell E, Bälter K. Feedback in web-based questionnaires as incentive to increase compliance in studies on lifestyle factors. *Public Health Nutr* 2012 Jun;15(6):982-988. [doi: [10.1017/S1368980011003041](https://doi.org/10.1017/S1368980011003041)] [Medline: [22122870](https://pubmed.ncbi.nlm.nih.gov/22122870/)]
27. Mathew M, Morrow JR, Frierson GM, Bain TM. Assessing digital literacy in web-based physical activity surveillance: the WIN study. *Am J Health Promot* 2011 Nov;26(2):90-95. [doi: [10.4278/ajhp.091001-QUAN-320](https://doi.org/10.4278/ajhp.091001-QUAN-320)] [Medline: [22040389](https://pubmed.ncbi.nlm.nih.gov/22040389/)]
28. Khosropour CM, Sullivan PS. Predictors of retention in an online follow-up study of men who have sex with men. *J Med Internet Res* 2011;13(3):e47 [FREE Full text] [doi: [10.2196/jmir.1717](https://doi.org/10.2196/jmir.1717)] [Medline: [21745792](https://pubmed.ncbi.nlm.nih.gov/21745792/)]
29. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. *J Med Internet Res* 2011;13(2):e38 [FREE Full text] [doi: [10.2196/jmir.1797](https://doi.org/10.2196/jmir.1797)] [Medline: [21571632](https://pubmed.ncbi.nlm.nih.gov/21571632/)]
30. Glasgow RE, Nelson CC, Kearney KA, Reid R, Ritzwoller DP, Strecher VJ, et al. Reach, engagement, and retention in an Internet-based weight loss program in a multi-site randomized controlled trial. *J Med Internet Res* 2007;9(2):e11 [FREE Full text] [doi: [10.2196/jmir.9.2.e11](https://doi.org/10.2196/jmir.9.2.e11)] [Medline: [17513282](https://pubmed.ncbi.nlm.nih.gov/17513282/)]
31. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. *J Med Internet Res* 2007;9(1):e1 [FREE Full text] [doi: [10.2196/jmir.9.1.e1](https://doi.org/10.2196/jmir.9.1.e1)] [Medline: [17478410](https://pubmed.ncbi.nlm.nih.gov/17478410/)]
32. Bull SS, Lloyd L, Rietmeijer C, McFarlane M. Recruitment and retention of an online sample for an HIV prevention intervention targeting men who have sex with men: the Smart Sex Quest Project. *AIDS Care* 2004 Nov;16(8):931-943. [doi: [10.1080/09540120412331292507](https://doi.org/10.1080/09540120412331292507)] [Medline: [15511725](https://pubmed.ncbi.nlm.nih.gov/15511725/)]
33. Bull SS, Vallejos D, Levine D, Ortiz C. Improving recruitment and retention for an online randomized controlled trial: experience from the Youthnet study. *AIDS Care* 2008 Sep;20(8):887-893. [doi: [10.1080/09540120701771697](https://doi.org/10.1080/09540120701771697)] [Medline: [18777217](https://pubmed.ncbi.nlm.nih.gov/18777217/)]
34. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009(3):MR000008. [doi: [10.1002/14651858.MR000008.pub4](https://doi.org/10.1002/14651858.MR000008.pub4)] [Medline: [19588449](https://pubmed.ncbi.nlm.nih.gov/19588449/)]

35. Edwards P, Roberts I, Clarke M, Diguseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009(3):MR0008-MR0008. [doi: [10.1002/14651858.MR000008.pub4](https://doi.org/10.1002/14651858.MR000008.pub4)] [Medline: [19588449](https://pubmed.ncbi.nlm.nih.gov/19588449/)]
36. Booker CL, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. *BMC Public Health* 2011;11:249 [FREE Full text] [doi: [10.1186/1471-2458-11-249](https://doi.org/10.1186/1471-2458-11-249)] [Medline: [21504610](https://pubmed.ncbi.nlm.nih.gov/21504610/)]
37. Robinson KA, Dennison CR, Wayman DM, Pronovost PJ, Needham DM. Systematic review identifies number of strategies important for retaining study participants. *J Clin Epidemiol* 2007 Aug;60(8):757-765 [FREE Full text] [doi: [10.1016/j.jclinepi.2006.11.023](https://doi.org/10.1016/j.jclinepi.2006.11.023)] [Medline: [17606170](https://pubmed.ncbi.nlm.nih.gov/17606170/)]
38. Wallace P, Murray E, McCambridge J, Khadjesari Z, White IR, Thompson SG, et al. On-line randomized controlled trial of an internet based psychologically enhanced intervention for people with hazardous alcohol consumption. *PLoS One* 2011;6(3):e14740 [FREE Full text] [doi: [10.1371/journal.pone.0014740](https://doi.org/10.1371/journal.pone.0014740)] [Medline: [21408060](https://pubmed.ncbi.nlm.nih.gov/21408060/)]
39. Dawson DA, Grant BF, Stinson FS, Zhou Y. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. *Alcohol Clin Exp Res* 2005 May;29(5):844-854. [Medline: [15897730](https://pubmed.ncbi.nlm.nih.gov/15897730/)]
40. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. *J Med Internet Res* 2009;11(2):e9 [FREE Full text] [doi: [10.2196/jmir.1052](https://doi.org/10.2196/jmir.1052)] [Medline: [19403465](https://pubmed.ncbi.nlm.nih.gov/19403465/)]
41. Khadjesari Z, Murray E, Kalaitzaki E, White IR, McCambridge J, Godfrey C, et al. Test-retest reliability of an online measure of past week alcohol consumption (the TOT-AL), and comparison with face-to-face interview. *Addict Behav* 2009 Apr;34(4):337-342. [doi: [10.1016/j.addbeh.2008.11.010](https://doi.org/10.1016/j.addbeh.2008.11.010)] [Medline: [19097705](https://pubmed.ncbi.nlm.nih.gov/19097705/)]
42. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001 Jul;33(5):337-343. [Medline: [11491192](https://pubmed.ncbi.nlm.nih.gov/11491192/)]
43. Linke S, McCambridge J, Khadjesari Z, Wallace P, Murray E. Development of a psychologically enhanced interactive online intervention for hazardous drinking. *Alcohol Alcohol* 2008;43(6):669-674 [FREE Full text] [doi: [10.1093/alcalc/agn066](https://doi.org/10.1093/alcalc/agn066)] [Medline: [18693217](https://pubmed.ncbi.nlm.nih.gov/18693217/)]
44. Cheung K, Oemar M, Oppe M, Rabin R. User Guide: Basic Information on How to Use EQ-5D. 2009 Mar. URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/User_Guide_v2_March_2009.pdf [accessed 2013-07-22] [WebCite Cache ID 6IInWOUcj]
45. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. New York: Wiley; 2000.
46. Khadjesari Z. *Use of the Internet for the delivery and evaluation of interventions aimed at reducing alcohol consumption* [thesis]. London: University College London; 2012.
47. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. *J Med Internet Res* 2006;8(3):e15 [FREE Full text] [doi: [10.2196/jmir.8.3.e15](https://doi.org/10.2196/jmir.8.3.e15)] [Medline: [16954125](https://pubmed.ncbi.nlm.nih.gov/16954125/)]

Abbreviations

- AUDIT-C:** Alcohol Use Disorders Identification Test (consumption questions)
DYD: Down Your Drink
OR: odds ratio
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 03.09.12; peer-reviewed by N Johnson, A Burdorf; comments to author 20.01.13; revised version received 04.04.13; accepted 08.05.13; published 30.08.13.

Please cite as:

Murray E, White IR, Varaganam M, Godfrey C, Khadjesari Z, McCambridge J
Attrition Revisited: Adherence and Retention in a Web-Based Alcohol Trial
J Med Internet Res 2013;15(8):e162
URL: <http://www.jmir.org/2013/8/e162/>
doi: [10.2196/jmir.2336](https://doi.org/10.2196/jmir.2336)
PMID: [23996958](https://pubmed.ncbi.nlm.nih.gov/23996958/)

©Elizabeth Murray, Ian R White, Mira Varaganam, Christine Godfrey, Zarnie Khadjesari, Jim McCambridge. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 30.08.2013. 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

Retrieving Clinical Evidence: A Comparison of PubMed and Google Scholar for Quick Clinical Searches

Salimah Z Shariff^{1,2}, BMath, PhD; Shayna AD Bejaimal¹, BMedSc; Jessica M Sontrop^{1,2}, PhD; Arthur V Iansavichus¹, MLIS; R Brian Haynes^{3,4}, MD, PhD; Matthew A Weir¹, MD; Amit X Garg^{1,2,3}, MD, PhD

¹Kidney Clinical Research Unit, Division of Nephrology, Western University, London, ON, Canada

²Department of Epidemiology and Biostatistics, Western University, London, ON, Canada

³McMaster University, Department of Clinical Epidemiology and Biostatistics, Hamilton, ON, Canada

⁴Department of Medicine, McMaster University, Hamilton, ON, Canada

Corresponding Author:

Salimah Z Shariff, BMath, PhD

Kidney Clinical Research Unit

Division of Nephrology

Western University

800 Commissioners Rd E. Rm ELL-108

London, ON, N6A 4G5

Canada

Phone: 1 519 685 8500 ext 56555

Fax: 1 519 685 8072

Email: salimah.shariff@lhsc.on.ca

Abstract

Background: Physicians frequently search PubMed for information to guide patient care. More recently, Google Scholar has gained popularity as another freely accessible bibliographic database.

Objective: To compare the performance of searches in PubMed and Google Scholar.

Methods: We surveyed nephrologists (kidney specialists) and provided each with a unique clinical question derived from 100 renal therapy systematic reviews. Each physician provided the search terms they would type into a bibliographic database to locate evidence to answer the clinical question. We executed each of these searches in PubMed and Google Scholar and compared results for the first 40 records retrieved (equivalent to 2 default search pages in PubMed). We evaluated the recall (proportion of relevant articles found) and precision (ratio of relevant to nonrelevant articles) of the searches performed in PubMed and Google Scholar. Primary studies included in the systematic reviews served as the reference standard for relevant articles. We further documented whether relevant articles were available as free full-texts.

Results: Compared with PubMed, the average search in Google Scholar retrieved twice as many relevant articles (PubMed: 11%; Google Scholar: 22%; $P<.001$). Precision was similar in both databases (PubMed: 6%; Google Scholar: 8%; $P=.07$). Google Scholar provided significantly greater access to free full-text publications (PubMed: 5%; Google Scholar: 14%; $P<.001$).

Conclusions: For quick clinical searches, Google Scholar returns twice as many relevant articles as PubMed and provides greater access to free full-text articles.

(*J Med Internet Res* 2013;15(8):e164) doi:[10.2196/jmir.2624](https://doi.org/10.2196/jmir.2624)

KEYWORDS

information dissemination/methods; information storage and retrieval; medical; library science; PubMed; Google Scholar; nephrology

Introduction

With the explosion of available health information, physicians increasingly rely on bibliographic databases for health information to guide the care of their patients. Unfortunately,

physicians face challenges when trying to find the information they need. They lack the time to develop efficient search strategies and often retrieve large numbers of nonrelevant articles [1-9]. Moreover, many bibliographic resources require

paid subscriptions, which further limit access to current best evidence.

Two online resources that are freely accessible around the world are PubMed and Google Scholar. PubMed remains the most widely used resource for medical literature [10]. More recently, Google Scholar has gained popularity as an alternative online bibliographic search resource [11-21]. Available search features in Google Scholar and PubMed are contrasted in Table 1. Whereas PubMed indexes only peer reviewed biomedical literature, Google Scholar also indexes articles, theses, books, abstracts, and court opinions from a variety of disciplines and sources including academic publishers, professional societies, online repositories, universities, and other websites [22]. While PubMed orders articles in roughly reverse chronological order, Google Scholar aims to order articles by relevance using a proprietary algorithm that weighs information from the full text

of each article, author, and journal information, and the number of times the article has been cited in other scholarly literature. Only a small fraction of the 21 million records in PubMed are available as free full-text publications via PubMed Central or specific journals. In contrast, due to its expanded search capabilities, Google Scholar may provide greater access to free full-text publications. To date, the utility of Google Scholar compared with PubMed for retrieving relevant primary literature to answer clinical questions has not been sufficiently tested.

In this study, we compare the ability of PubMed and Google Scholar to retrieve relevant renal literature for searches created by nephrologists to address questions of renal therapy. Renal literature is dispersed across more than 400 multidisciplinary journals and is difficult for nephrologists to track down; thus, this discipline provides an ideal model for this type of evaluation [23].

Table 1. Search features available in PubMed and Google Scholar.

Feature	PubMed	Google Scholar
Searching		
Allows use of Boolean terms (AND, OR, NOT)	Yes	Yes
Provides search limits (eg, age, publication type, date)	Yes (extensive)	Yes (very limited)
Provides use of search filters that limit search results to a specific clinical study category or subject matter (eg, Clinical Queries, Topic-Specific Queries)	Yes	No
Allows use of truncation (inclusion of multiple beginnings or endings achieved by typing in an asterisk "*" in PubMed – eg, cardio*)	Yes	No (automatically searches for variances in words)
Allows use of controlled vocabulary (eg, MeSH terminology)	Yes	No
Provides spell checking for search terms	Yes	Yes
Stores search history	Yes	No
Sorts results by relevance	No	Yes
Access to articles		
Indicates whether articles are available as free full-texts	Yes	Yes
Allows linking to institutions for subscription access (eg, link to university library)	Yes	Yes
Other services		
Allows article citations to be imported into bibliography managers (eg, Reference Manager)	Yes (can import multiple selected citations)	Yes (can import only one citation at time)
Tracks the number of times articles are cited by other publications	Yes (only to journals in PubMed Central)	Yes
When searching, algorithm also searches the full-text of publications	No	Yes
Provides email alerts for prespecified searches	Yes	Yes (introduced in 2010)
Allows users to view related articles for an article of interest	Yes	Yes
Source lists all journals and their publication dates that are included in their data holdings.	Yes	No

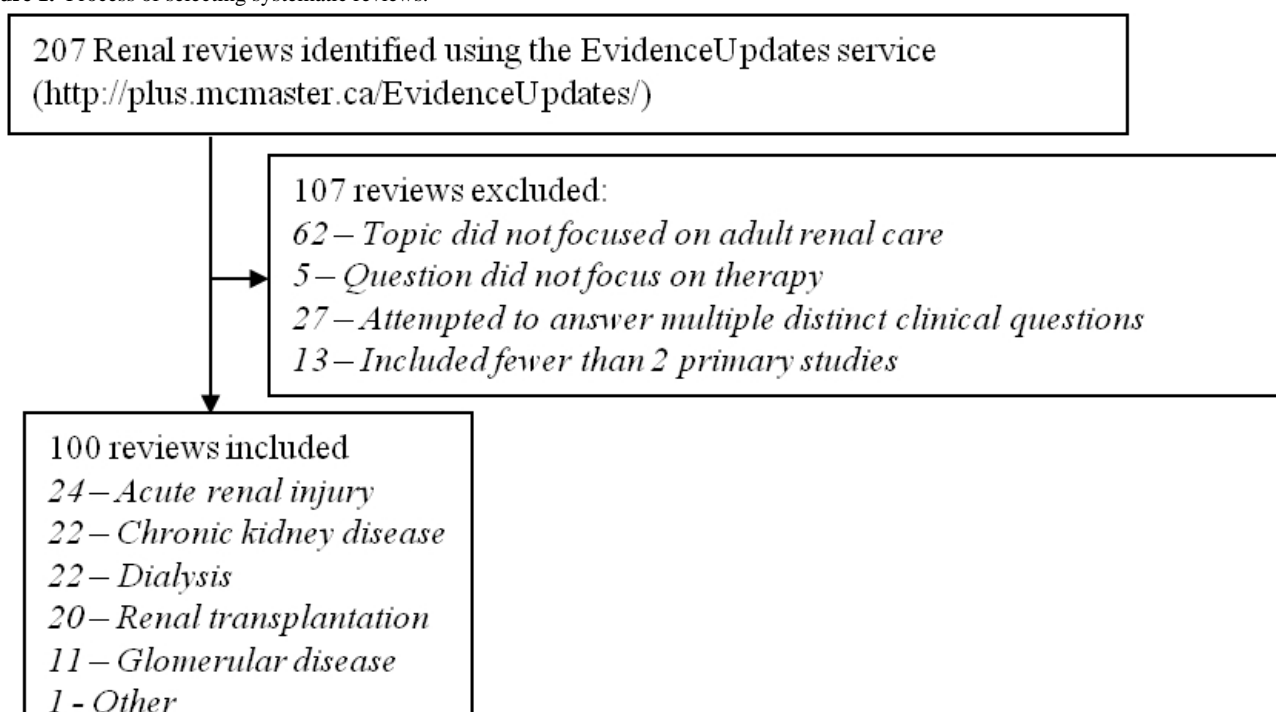
Methods

Clinical Questions and Articles of Interest

We derived 100 clinical questions from the objectives statements of 100 high-quality systematic reviews on renal therapy published between 2001 and 2009. We identified the systematic reviews from the EvidenceUpdates service in November 2009, by selecting the option to view all reviews for the discipline of nephrology; our search yielded 207 systematic reviews. This service prescreens and identifies systematic reviews and meta-analyses that meet strict methodological criteria and have a high potential for clinical relevance [24,25]. Two nephrologists used a previously developed checklist to independently confirm that each review targeted a single clinical question relevant to adult nephrology care (kappa=0.98) [26] and included at least 2 primary studies. Discrepancies were resolved through

discussion; 100 reviews met the inclusion criteria (see Figure 1 for the process of selecting reviews). We transformed the objectives statement from each review into a clinical question (see Multimedia Appendix 1 for a sample of the questions posed and search queries received). For example, the objective of one review was “to assess the effectiveness of normal saline versus sodium bicarbonate for prevention of contrast-induced nephropathy”. We transformed this statement into the clinical question: “How effective is normal saline versus sodium bicarbonate for the prevention of contrast-induced nephropathy?”[27]. We extracted the citations to the primary studies referenced in each review that met the eligibility criteria for inclusion. These citations acted as a set of relevant articles for the corresponding clinical question (also referred to as the “reference standard”). The reviews cited a mean of 19 articles, totaling 1574 unique citations across all reviews.

Figure 1. Process of selecting systematic reviews.



Data Collection and Measurements

We surveyed a simple random sample of nephrologists practicing in Canada (response rate 75%). Survey details are available elsewhere [28,29]. Briefly, we asked nephrologists about their information-gathering practices. In addition, we provided each nephrologist with a unique, randomly selected therapy-focused clinical question generated from a renal systematic review. The nephrologists provided the search terms they would type into a bibliographic resource to retrieve relevant studies to address the clinical question (known as a “search query”). The survey was designed and administered using the Dillman tailored design method [30]. The sampling frame consisted of nephrologists practicing in Canada and included both academic (practicing in a center with a fellowship training program) and community-based nephrologists. Nephrologists were selected from the sampling frame using a random number generator; one nephrologist was selected at a time and randomly

assigned a clinical question. Once a selected nephrologist was deemed a nonresponder, the same clinical question was assigned to another nephrologist. In addition, upon receipt of a completed survey, if a respondent did not provide a search query to the clinical question, the same survey was re-administered to a new participant. Survey administration continued until 1 valid search query for each of the 100 questions was received. Overall, 115 survey responses were received and 15 were excluded from further analysis because of missing or illegible search queries (n=8) or because the survey was received after a physician was deemed a nonresponder (n=7).

To compare the performance of PubMed and Google Scholar for use by practicing physicians, we executed each physician-provided search query in PubMed and Google Scholar using all default settings. Occasionally, physicians provided misspelled search terms, acronyms, or other discrepancies. To address this, the syntax of the search was modified slightly using

prespecified rules (listed in [Multimedia Appendix 2](#)). This was done in duplicate and differences were resolved by consensus. All searches were conducted between May and July 2010. We restricted each search to the search dates provided in the methods section of each systematic review. For each search result, we calculated the total number of records retrieved, the number of relevant articles retrieved, and the position of the relevant records in the search results. For each relevant article, we followed all links for full-text access and documented whether the full-text version could be retrieved for free. We did not use any pay-for-privilege accesses. To ensure that we did not inadvertently make use of our institution's licensing when searching, all searches were conducted on a computer with Internet access provided by a local service provider and not our institution. We tested and validated our search methodology in a pilot phase. Two assessors (graduate students with expertise in computer science and biomedical science) independently conducted 10 searches in PubMed and Google Scholar and achieved a percent agreement of 99%.

Content Coverage

To assess the potential for bias due to the absence of articles in one source over the other, we evaluated the content coverage for each database. A content coverage analysis determines whether pertinent literature is contained within a specific bibliographic database [31]. There are two potential reasons for not finding an important article when searching a database such as PubMed: either the article of interest is not included in the content holdings of the database (referred to as a lack of content coverage), or the article is present, but the search mechanism fails to retrieve it when a search phrase is typed into the database. To determine content coverage, we searched for each

primary article using advanced search strategies as outlined in other coverage studies [32,33]. This involved various combinations of the manuscript's title (both English and non-English), the authors' names, journal title, page numbers, and the year published. We selected all links to candidate matches to confirm a match. In Google Scholar, the option to view all versions for a candidate article was always selected and all links were attempted. If a primary article was not found in one of the resources, further searches were performed by another rater to confirm its absence. We previously published a more comprehensive content coverage analysis of renal literature that applied the same methods [34].

General Statistical Analytic Strategy and Sample Size

Primary Analysis

The two most prominent performance metrics of searching are recall and precision (Table 2). Results from our survey indicated that 80% of nephrologists do not review beyond 40 search results, which is the equivalent of 2 default search pages in PubMed [28]. Thus, for the primary analysis, we calculated the recall and precision for the first 40 retrieved records in each search. We used a 2-sided paired *t* test to compare search outcomes between PubMed and Google Scholar. To reduce the risk of type I error, we used a conservative *P* value of .025 to interpret significance for all comparisons. We used SAS, Version 9.2 for all statistical analyses.

Secondary Analysis

We repeated the calculation for recall while only considering relevant articles that are freely accessible. For each physician-generated search, we also calculated the recall and precision for all retrieved records (not just the first 40).

Table 2. Formulas for calculating search recall^a and precision^b.

Search in PubMed or Google Scholar for a clinical question	Relevant articles ^c	Nonrelevant articles
Articles found	TP	FP
Articles not found	FN	TN

^aSearch recall $TP/(TP + FN)$: the number of relevant articles found as a proportion of the total number of relevant articles.

^bSearch precision $TP/(TP + FP)$ (also referred to as the positive predictive value in diagnostic test terminology): the number of relevant articles found as a proportion of the total number of articles found.

^cFor each search, the set of relevant articles were the collection primary studies included in the original systematic review from which the clinical question was derived.

Results

Nephrologist and Search Characteristics

Participating nephrologists were an average of 48 years old and had practiced nephrology for an average of 15 years. All respondents had used an online resource to guide the treatment of their patients in the previous year. Approximately 90% used PubMed to search, while 40% used Google Scholar; 32% indicated using both bibliographic resources. Searches provided by the nephrologists contained an average of three concept terms, with each term embodying a single concept, for example, myocardial infarction. Forty-eight percent of nephrologists used Boolean terms such as AND, OR, and NOT in their searches. Seven percent of searches included advanced search features

such as search limits, search filters, and truncation (inclusion of multiple endings achieved by typing in an asterisk "*" in PubMed, eg, nephro*). No substantive differences were observed in searches provided by older versus younger nephrologists, males versus females, or by those practicing in an academic versus community setting.

Content Coverage

PubMed and Google Scholar contained similar proportions of tested articles in their database holdings: each contained 78% of the 1574 unique citations collected. Google Scholar contained an additional 5% of the articles not included in PubMed and PubMed contained an additional 2% of the articles not included in Google Scholar; 15% of the articles were missing in both sources.

Primary Analysis

Google Scholar retrieved twice as many relevant articles as PubMed within the first 40 records (average recall: 21.9% vs 10.9%; [Table 3](#)). Precision was similar in the two databases. When we considered both metrics together, Google Scholar demonstrated better recall and similar precision in 77% of searches.

Secondary Analysis

Google Scholar retrieved three times as many relevant articles with free full-text access compared with PubMed (average

recall: 15% vs 5%; $P < 0.001$; [Table 3](#)). When examining all records (not just the first 40 records), PubMed and Google Scholar retrieved a similar number of relevant articles, although Google Scholar continued to provide increased free full-text access to publications. Overall, searches in Google Scholar retrieved more records per search compared with PubMed (median: 1000 records vs 148 records, respectively). This resulted in lower search precision in Google Scholar when all retrieved articles were considered.

Table 3. Recall and precision of physician searches tested in PubMed and Google Scholar (within the first 40 citations, PubMed found no relevant citations for 54% of the searches and Google Scholar found no relevant citations for 21% of the searches).

Measure ^a	Within first 40 citations		All citations	
	Mean, % ^b (SD ^c)	<i>P</i> value ^d	Mean, % ^b (SD ^c)	<i>P</i> value ^d
Recall		<.001		.10
PubMed	10.9 (20)		38.0 (33)	
Google Scholar	21.9 (24)		43.2 (29)	
Free full-text recall		<.001		<.001
PubMed	4.7 (11)		16.4 (20)	
Google Scholar	14.6 (20)		25.1 (23)	
Precision		.07		<.001
PubMed	5.6 (11)		6.0 (11)	
Google Scholar	7.6 (7)		0.8 (0.8)	

^aFormulas for measures: (1) Recall: (number of relevant articles retrieved) / (total number of relevant articles available); (2) Free full-text recall: (number of relevant articles retrieved available for free full-text viewing) / (total number of relevant articles available), and (3) Precision: (number of relevant articles retrieved) / (total number of citations retrieved).

^bValues represent mean of results from 100 searches.

^cSD=Standard deviation.

^d*P* values compare PubMed to Google Scholar using a paired *t* test; significance values remained similar when using the nonparametric Wilcoxon signed rank test.

Discussion

Principal Findings

Nephrologists increasingly rely on online bibliographic databases to guide the care of their patients. Because most nephrologists view fewer than 40 search results, important literature will be missed if not contained within this narrow window [28]. For our primary objective, we compared the ability of PubMed and Google Scholar to retrieve relevant renal literature within the first 40 records. We found that Google Scholar retrieved twice as many relevant articles as PubMed—and three times as many relevant articles with free full-text access. These results are not attributable to differences in content coverage, as 78% of the tested articles were available in both databases [34]. Instead, the improved performance of Google Scholar may result from its method of ranking results based on relevance. However, when considering all search results (not just the first 40 records), recall was similar between Google Scholar and PubMed, while precision favored PubMed.

While many academics see the value in Google Scholar, its uptake has been slow among physicians [11-21,35-40]. Unlike

Google Scholar, PubMed provides indexed content that is directly relevant to physicians, including clinical controlled vocabulary (MeSH [medical subject headings]), search limits (such as limiting articles by age or study type), and access to discipline-specific and methods search filters [24,41-43]. These advanced features have the potential to reduce the number of nonrelevant articles that are retrieved. However, only 7% of respondents used these features in their searches for this study. While 77% of nephrologists reported previous use of search limits, only 37% used controlled vocabularies, and only 20% used filters such as the Clinical Queries feature in PubMed [28,29]. Whereas PubMed searches retrieve published literature from biomedical journals, Google Scholar searches retrieve both published and unpublished literature from a range of disciplines. This may explain the greater overall number of records found per search (median of 1000 for Google Scholar and 148 for PubMed).

Google Scholar provided significantly greater access to free full-text articles. This is notable given that physicians and institutions in developing nations may lack the resources needed to maintain subscriptions to journals. Even in developed countries, the burden of paying for knowledge is felt. Some

academic databases and journals have raised their fees for university subscriptions by up to 400%. This prompted one Canadian university library to cancel subscription access to the Web of Science bibliographic database, citing “a challenging fiscal climate” as a primary reason [44-46].

Our findings are consistent with those of previous studies [12,14,15,20,21]. In preliminary testing within targeted areas of respiratory care, sarcoma, pharmacotherapy, and family medicine, Google Scholar provided better comprehensiveness (recall) but worse efficiency (precision) compared with PubMed. Similar results were seen in our study when we considered all records that were retrieved and not just the first 40. However, previous studies tested only a small number of searches (range: 1-22), compared with the 100 searches in the current study. In addition, the search queries used in previous studies were created and tested by researchers in idealized settings, which may not generalize as well to searches generated by physicians in busy clinical settings.

We followed recommendations of search database evaluations from the field of information retrieval and designed our study to improve on limitations of previous studies [47,48]. To ensure that the clinical questions tested were relevant to practicing nephrologists, we gathered questions using renal systematic reviews that targeted questions in patient care where uncertainty exists. To ensure that all articles in the review were relevant for the clinical question, we selected systematic reviews that specified only one objective. Finally, to maximize external validity, we used a survey to obtain searches created by practicing nephrologists. Our survey achieved a response rate of 75% with responses from both newer and more seasoned nephrologists practicing in both community and academic settings [28].

Limitations

Our study has some limitations. We did not directly observe the nephrologists as they searched. There may be a discrepancy

between what search terms busy clinicians report in a survey and what they actually type in practice [37]. As recommended, we used primary studies included in high-quality systematic reviews to define relevance [14,20,49-54]. By using this method, we were unable to include other articles that some physicians may find relevant (eg, studies of lower methodological quality, narrative reviews, case reports, commentaries). However, our approach engages widely accepted principles of the hierarchy of evidence to identify the most important primary articles to retrieve in a search. For reasons of feasibility, our study focused on questions of therapy. As more systematic reviews for diagnosis, prognosis, and etiology are published, we will be able to expand this study to test searches for these types of studies as well. Our study tests the searches provided by the physicians, which are likely initial searches; however, in practice, an unsatisfactory search may be attempted again or modified based on the results obtained. Yet, our results indicate that Google Scholar can improve on the nephrologists' initial search, which can save valuable clinical time. Given the nature of the survey, we are uncertain about how many steps nephrologists take to refine their search and future research should explore this. Although Google Scholar retrieved twice as many relevant articles as PubMed (in the first 40 citations), 80% of relevant articles were not retrieved by either source. Improved methods to efficiently retrieve relevant articles for clinical decision making requires further development and testing. Future research might also evaluate the effectiveness of strategies that apply relevance-based ranking to PubMed results on physician searches [55].

Conclusions

In conclusion, for quick clinical searches, Google Scholar returns twice as many relevant articles as PubMed and provides greater access to free full-texts. Improved searching by clinicians has the potential to enhance the transfer of research into practice and improve patient care. Future studies should confirm these results for other medical disciplines.

Acknowledgments

Funding for this study came from “The Physicians’ Services Inc. Foundation”. Dr Shariff was supported by a Canadian Institutes of Health Research (CIHR) Doctoral Research Award and the Schulich Scholarship for Medical Research from Western University. Dr Garg was supported by a Clinician Scientist Award from the CIHR.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample of systematic reviews selected and search queries received by respondents.

[PDF File (Adobe PDF File), 185KB - [jmir_v15i8e164_app1.pdf](#)]

Multimedia Appendix 2

Rules used for syntactically improving physician searches in PubMed and Google Scholar.

[PDF File (Adobe PDF File), 257KB - [jmir_v15i8e164_app2.pdf](#)]

References

1. Ely JW, Osheroff JA, Ebell MH, Chambliss ML, Vinson DC, Stevermer JJ, et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ* 2002 Mar 23;324(7339):710 [FREE Full text] [Medline: [11909789](#)]
2. Chambliss ML, Conley J. Answering clinical questions. *J Fam Pract* 1996 Aug;43(2):140-144. [Medline: [8708623](#)]
3. Norlin C, Sharp AL, Firth SD. Unanswered questions prompted during pediatric primary care visits. *Ambul Pediatr* 2007;7(5):396-400. [doi: [10.1016/j.ambp.2007.05.008](#)] [Medline: [17870649](#)]
4. Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. *Health Info Libr J* 2007 Jun;24(2):78-94. [doi: [10.1111/j.1471-1842.2007.00713.x](#)] [Medline: [17584211](#)]
5. Gorman P. Does the medical literature contain the evidence to answer the questions of primary care physicians? Preliminary findings of a study. *Proc Annu Symp Comput Appl Med Care* 1993:571-575 [FREE Full text] [Medline: [8130538](#)]
6. Gorman PN, Helfand M. Information seeking in primary care: how physicians choose which clinical questions to pursue and which to leave unanswered. *Med Decis Making* 1995 Apr;15(2):113-119. [Medline: [7783571](#)]
7. Ely JW, Osheroff JA, Ebell MH, Bergus GR, Levy BT, Chambliss ML, et al. Analysis of questions asked by family doctors regarding patient care. *BMJ* 1999 Aug 7;319(7206):358-361 [FREE Full text] [Medline: [10435959](#)]
8. González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, et al. Information needs and information-seeking behavior of primary care physicians. *Ann Fam Med* 2007 Jul;5(4):345-352 [FREE Full text] [doi: [10.1370/afm.681](#)] [Medline: [17664501](#)]
9. Ely JW, Osheroff JA, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc* 2005 Mar;12(2):217-224 [FREE Full text] [doi: [10.1197/jamia.M1608](#)] [Medline: [15561792](#)]
10. National Library of Medicine (US). Key MEDLINE Indicators. URL: http://www.nlm.nih.gov/bsd/bsd_key.html [accessed 2013-03-17] [WebCite Cache ID 6FCUenHlf]
11. Younger P. Using google scholar to conduct a literature search. *Nurs Stand* 2010;24(45):40-6; quiz 48. [Medline: [20701052](#)]
12. Mastrangelo G, Fadda E, Rossi CR, Zamprogno E, Buja A, Cegolon L. Literature search on risk factors for sarcoma: PubMed and Google Scholar may be complementary sources. *BMC Res Notes* 2010;3:131 [FREE Full text] [doi: [10.1186/1756-0500-3-131](#)] [Medline: [20459746](#)]
13. Kulkarni AV, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *JAMA* 2009 Sep 9;302(10):1092-1096. [doi: [10.1001/jama.2009.1307](#)] [Medline: [19738094](#)]
14. Freeman MK, Lauderdale SA, Kendrach MG, Woolley TW. Google Scholar versus PubMed in locating primary literature to answer drug-related questions. *Ann Pharmacother* 2009 Mar;43(3):478-484. [doi: [10.1345/aph.1L223](#)] [Medline: [19261965](#)]
15. Shultz M. Comparing test searches in PubMed and Google Scholar. *J Med Libr Assoc* 2007 Oct;95(4):442-445 [FREE Full text] [doi: [10.3163/1536-5050.95.4.442](#)] [Medline: [17971893](#)]
16. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB J* 2008 Feb;22(2):338-342 [FREE Full text] [doi: [10.1096/fj.07-9492LSF](#)] [Medline: [17884971](#)]
17. Henderson J. Google Scholar: A source for clinicians? *CMAJ* 2005 Jun 7;172(12):1549-1550 [FREE Full text] [doi: [10.1503/cmaj.050404](#)] [Medline: [15939908](#)]
18. Giustini D, Barsky E. A look at Google Scholar, PubMed, and Scirus: comparisons and recommendations. *Journal of the Canadian Health Libraries Association* 2005;26(3):85-89.
19. Vine R. Google Scholar electronic resources review. *Journal of the Medical Library Association* 2006;94(1):---
20. Anders ME, Evans DP. Comparison of PubMed and Google Scholar literature searches. *Respir Care* 2010 May;55(5):578-583 [FREE Full text] [Medline: [20420728](#)]
21. Nourbakhsh E, Nugent R, Wang H, Cevik C, Nugent K. Medical literature searches: a comparison of PubMed and Google Scholar. *Health Info Libr J* 2012 Sep;29(3):214-222. [doi: [10.1111/j.1471-1842.2012.00992.x](#)] [Medline: [22925384](#)]
22. Google Scholar Beta: About Google Scholar. URL: <http://scholar.google.com/scholar/about.html> [accessed 2013-03-17] [WebCite Cache ID 6FCUrUKMS]
23. Garg AX, Iansavichus AV, Kastner M, Walters LA, Wilczynski N, McKibbin KA, et al. Lost in publication: Half of all renal practice evidence is published in non-renal journals. *Kidney Int* 2006 Dec;70(11):1995-2005. [doi: [10.1038/sj.ki.5001896](#)] [Medline: [17035946](#)]
24. Haynes RB, McKibbin KA, Wilczynski NL, Walter SD, Werre SR, Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005 May 21;330(7501):1179 [FREE Full text] [doi: [10.1136/bmj.38446.498542.8F](#)] [Medline: [15894554](#)]
25. Haynes RB. *bmjupdates+*, a new free service for evidence-based clinical practice. *Evid Based Nurs* 2005 Apr;8(2):39 [FREE Full text] [Medline: [15830413](#)]
26. Garg AX, Iansavichus AV, Wilczynski NL, Kastner M, Baier LA, Shariff SZ, et al. Filtering Medline for a clinical discipline: diagnostic test assessment framework. *BMJ* 2009;339:b3435 [FREE Full text] [Medline: [19767336](#)]

27. Meier P, Ko DT, Tamura A, Tamhane U, Gurm HS. Sodium bicarbonate-based hydration prevents contrast-induced nephropathy: a meta-analysis. *BMC Med* 2009;7:23 [FREE Full text] [doi: [10.1186/1741-7015-7-23](https://doi.org/10.1186/1741-7015-7-23)] [Medline: [19439062](https://pubmed.ncbi.nlm.nih.gov/19439062/)]
28. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Weir MA, Haynes RB, et al. Searching for medical information online: a survey of Canadian nephrologists. *J Nephrol* 2011;24(6):723-732. [doi: [10.5301/JN.2011.6373](https://doi.org/10.5301/JN.2011.6373)] [Medline: [21360475](https://pubmed.ncbi.nlm.nih.gov/21360475/)]
29. Shariff SZ, Sontrop JM, Haynes RB, Iansavichus AV, McKibbin KA, Wilczynski NL, et al. Impact of PubMed search filters on the retrieval of evidence by physicians. *CMAJ* 2012 Feb 21;184(3):E184-E190 [FREE Full text] [doi: [10.1503/cmaj.101661](https://doi.org/10.1503/cmaj.101661)] [Medline: [22249990](https://pubmed.ncbi.nlm.nih.gov/22249990/)]
30. Dillman DA. *Mail and internet surveys: The tailored design method*. New Jersey: John Wiley & Sons Inc; 2007.
31. Lancaster FW. The evaluation of published indexes and abstract journals: criteria and possible procedures. *Bull Med Libr Assoc* 1971 Jul;59(3):479-494 [FREE Full text] [Medline: [5146770](https://pubmed.ncbi.nlm.nih.gov/5146770/)]
32. Christianson M. Ecology articles in Google Scholar: Levels of Access to Articles in Core Journals. *Issues in Science and Technology Librarianship* 2007;49:---. [doi: [10.5062/F4MS3QPD](https://doi.org/10.5062/F4MS3QPD)]
33. Neuhaus C, Neuhaus E, Asher A, Wrede C. The Depth and Breadth of Google Scholar: An Empirical Study. *portal: Libraries and the Academy* 2006;6(2):127-141. [doi: [10.1353/PLA](https://doi.org/10.1353/PLA)]
34. Shariff SZ, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, Gandhi S, et al. Availability of renal literature in six bibliographic databases. *Clin Kidney J* 2012 Dec;5(6):610-617 [FREE Full text] [doi: [10.1093/ckj/sfs152](https://doi.org/10.1093/ckj/sfs152)] [Medline: [23185693](https://pubmed.ncbi.nlm.nih.gov/23185693/)]
35. De Leo G, LeRouge C, Ceriani C, Niederman F. Websites most frequently used by physician for gathering medical information. *AMIA Annu Symp Proc* 2006:902 [FREE Full text] [Medline: [17238521](https://pubmed.ncbi.nlm.nih.gov/17238521/)]
36. Somal K, Lam WC, Tam E. Computer and internet use by ophthalmologists and trainees in an academic centre. *Can J Ophthalmol* 2009 Jun;44(3):265-268. [doi: [10.3129/i09-057](https://doi.org/10.3129/i09-057)] [Medline: [19491979](https://pubmed.ncbi.nlm.nih.gov/19491979/)]
37. Chiu YW, Weng YH, Lo HL, Ting HW, Hsu CC, Shih YH, et al. Physicians' characteristics in the usage of online database: a representative nationwide survey of regional hospitals in Taiwan. *Inform Health Soc Care* 2009 Sep;34(3):127-135. [doi: [10.1080/17538150903102372](https://doi.org/10.1080/17538150903102372)] [Medline: [19670003](https://pubmed.ncbi.nlm.nih.gov/19670003/)]
38. Kitchin DR, Applegate KE. Learning radiology a survey investigating radiology resident use of textbooks, journals, and the internet. *Acad Radiol* 2007 Sep;14(9):1113-1120. [doi: [10.1016/j.acra.2007.06.002](https://doi.org/10.1016/j.acra.2007.06.002)] [Medline: [17707320](https://pubmed.ncbi.nlm.nih.gov/17707320/)]
39. Hider PN, Griffin G, Walker M, Coughlan E. The information-seeking behavior of clinical staff in a large health care organization. *J Med Libr Assoc* 2009 Jan;97(1):47-50 [FREE Full text] [doi: [10.3163/1536-5050.97.1.009](https://doi.org/10.3163/1536-5050.97.1.009)] [Medline: [19159006](https://pubmed.ncbi.nlm.nih.gov/19159006/)]
40. Giustini D. How Google is changing medicine. *BMJ* 2005 Dec 24;331(7531):1487-1488 [FREE Full text] [doi: [10.1136/bmj.331.7531.1487](https://doi.org/10.1136/bmj.331.7531.1487)] [Medline: [16373722](https://pubmed.ncbi.nlm.nih.gov/16373722/)]
41. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ* 2004 May 1;328(7447):1040 [FREE Full text] [doi: [10.1136/bmj.38068.557998.EE](https://doi.org/10.1136/bmj.38068.557998.EE)] [Medline: [15073027](https://pubmed.ncbi.nlm.nih.gov/15073027/)]
42. Wilczynski NL, Haynes RB, Hedges Team. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. *BMC Med* 2004 Jun 9;2:23 [FREE Full text] [doi: [10.1186/1741-7015-2-23](https://doi.org/10.1186/1741-7015-2-23)] [Medline: [15189561](https://pubmed.ncbi.nlm.nih.gov/15189561/)]
43. Wilczynski NL, Haynes RB, Hedges Team. Developing optimal search strategies for detecting clinically sound causation studies in MEDLINE. *AMIA Annu Symp Proc* 2003:719-723 [FREE Full text] [Medline: [14728267](https://pubmed.ncbi.nlm.nih.gov/14728267/)]
44. Howard J. Canadian University Hopes to Lead Fight Against High Subscription Prices. 2010. URL: <https://chronicle.com/article/Canadian-University-Hopes-to-66095/> [accessed 2013-03-17] [WebCite Cache ID 6FCVHnUW0]
45. Taylor L. Canadian librarian leads worldwide digital revolt for free knowledge. 2010. URL: <http://www.thestar.com/life/2010/08/10/canadian-librarian-leads-worldwide-digital-revolt-for-free-knowledge.html> [accessed 2013-03-17] [WebCite Cache ID 6FCWjRIXK]
46. Howard J. University of California Tries Just Saying No to Rising Journal Costs. 2010. URL: <http://chronicle.com/article/U-of-California-Tries-Just/65823/> [accessed 2013-03-17] [WebCite Cache ID 6FCWe0rCE]
47. Gordon M, Pathak P. Finding information on the World Wide Web: the retrieval effectiveness of search engines. *Information Processing & Management* 1999 Mar;35(2):141-180. [doi: [10.1016/S0306-4573\(98\)00041-7](https://doi.org/10.1016/S0306-4573(98)00041-7)]
48. Hersh WR. *Information retrieval: a health and biomedical perspective*. New York: Springer Verlag; 2008.
49. Slobogean GP, Verma A, Giustini D, Slobogean BL, Mulpuri K. MEDLINE, EMBASE, and Cochrane index most primary studies but not abstracts included in orthopedic meta-analyses. *J Clin Epidemiol* 2009 Dec;62(12):1261-1267. [doi: [10.1016/j.jclinepi.2009.01.013](https://doi.org/10.1016/j.jclinepi.2009.01.013)] [Medline: [19364634](https://pubmed.ncbi.nlm.nih.gov/19364634/)]
50. Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *J Clin Epidemiol* 2005 Jun;58(6):560-567. [doi: [10.1016/j.jclinepi.2004.11.018](https://doi.org/10.1016/j.jclinepi.2004.11.018)] [Medline: [15878469](https://pubmed.ncbi.nlm.nih.gov/15878469/)]
51. Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. *Cochrane Database Syst Rev* 2007(2):MR000001. [doi: [10.1002/14651858.MR000001.pub2](https://doi.org/10.1002/14651858.MR000001.pub2)] [Medline: [17443625](https://pubmed.ncbi.nlm.nih.gov/17443625/)]

52. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Med Res Methodol* 2006;6:33 [[FREE Full text](#)] [doi: [10.1186/1471-2288-6-33](https://doi.org/10.1186/1471-2288-6-33)] [Medline: [16848895](https://pubmed.ncbi.nlm.nih.gov/16848895/)]
53. Yousefi-Nooraie R, Irani S, Mortaz-Hedjri S, Shakiba B. Comparison of the efficacy of three PubMed search filters in finding randomized controlled trials to answer clinical questions. *J Eval Clin Pract* 2010 Sep 16:-. [doi: [10.1111/j.1365-2753.2010.01554.x](https://doi.org/10.1111/j.1365-2753.2010.01554.x)] [Medline: [20846321](https://pubmed.ncbi.nlm.nih.gov/20846321/)]
54. Agoritsas T, Merglen A, Courvoisier DS, Combescure C, Garin N, Perrier A, et al. Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. *J Med Internet Res* 2012;14(3):e85 [[FREE Full text](#)] [doi: [10.2196/jmir.2021](https://doi.org/10.2196/jmir.2021)] [Medline: [22693047](https://pubmed.ncbi.nlm.nih.gov/22693047/)]
55. Lu Z, Kim W, Wilbur WJ. Evaluating relevance ranking strategies for MEDLINE retrieval. *J Am Med Inform Assoc* 2009 Feb;16(1):32-36 [[FREE Full text](#)] [doi: [10.1197/jamia.M2935](https://doi.org/10.1197/jamia.M2935)] [Medline: [18952932](https://pubmed.ncbi.nlm.nih.gov/18952932/)]

Edited by G Eysenbach; submitted 18.03.13; peer-reviewed by N Allen, D Perez-Rey, A Manconi; comments to author 04.05.13; revised version received 16.05.13; accepted 11.06.13; published 15.08.13.

Please cite as:

Shariff SZ, Bejaimal SAD, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, Garg AX
Retrieving Clinical Evidence: A Comparison of PubMed and Google Scholar for Quick Clinical Searches
J Med Internet Res 2013;15(8):e164
URL: <http://www.jmir.org/2013/8/e164/>
doi: [10.2196/jmir.2624](https://doi.org/10.2196/jmir.2624)
PMID: [23948488](https://pubmed.ncbi.nlm.nih.gov/23948488/)

©Salimah Z Shariff, Shayna AD Bejaimal, Jessica M Sontrop, Arthur V Iansavichus, R Brian Haynes, Matthew A Weir, Amit X Garg. Originally published in the *Journal of Medical Internet Research* (<http://www.jmir.org>), 15.08.2013. 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 Method for the Design and Development of Medical or Health Care Information Websites to Optimize Search Engine Results Page Rankings on Google

Suzanne Dunne¹, BSc (Hons), MSc; Niamh Maria Cummins¹, BSc, MSc, PhD; Ailish Hannigan^{1,2}, BSc, PhD; Bill Shannon¹, FRCGP, MICGP, MD; Colum Dunne^{1,2}, BSc (Hons), MBA, PhD; Walter Cullen^{1,2}, MB, MICGP, MD

¹Graduate Entry Medical School, University of Limerick, Limerick, Ireland

²Centre for Interventions in Infection, Inflammation and Immunity (4i), Graduate Entry Medical School, University of Limerick, Limerick, Ireland

Corresponding Author:

Suzanne Dunne, BSc (Hons), MSc
Graduate Entry Medical School
University of Limerick
Limerick
Ireland
Phone: 353 (0)868560296
Fax: 353 (0)61 233778
Email: suzanne.dunne@ul.ie

Abstract

Background: The Internet is a widely used source of information for patients searching for medical/health care information. While many studies have assessed existing medical/health care information on the Internet, relatively few have examined methods for design and delivery of such websites, particularly those aimed at the general public.

Objective: This study describes a method of evaluating material for new medical/health care websites, or for assessing those already in existence, which is correlated with higher rankings on Google's Search Engine Results Pages (SERPs).

Methods: A website quality assessment (WQA) tool was developed using criteria related to the quality of the information to be contained in the website in addition to an assessment of the readability of the text. This was retrospectively applied to assess existing websites that provide information about generic medicines. The reproducibility of the WQA tool and its predictive validity were assessed in this study.

Results: The WQA tool demonstrated very high reproducibility (intraclass correlation coefficient=0.95) between 2 independent users. A moderate to strong correlation was found between WQA scores and rankings on Google SERPs. Analogous correlations were seen between rankings and readability of websites as determined by Flesch Reading Ease and Flesch-Kincaid Grade Level scores.

Conclusions: The use of the WQA tool developed in this study is recommended as part of the design phase of a medical or health care information provision website, along with assessment of readability of the material to be used. This may ensure that the website performs better on Google searches. The tool can also be used retrospectively to make improvements to existing websites, thus, potentially enabling better Google search result positions without incurring the costs associated with Search Engine Optimization (SEO) professionals or paid promotion.

(*J Med Internet Res* 2013;15(8):e183) doi:[10.2196/jmir.2632](https://doi.org/10.2196/jmir.2632)

KEYWORDS

health care information; patient education; Google; Internet; medical informatics; generic drugs; website development; quality assessment

Introduction

A multitude of studies have assessed the use, quality, and availability of medical/health care information on the Internet

in areas as diverse as inflammatory bowel disease [1], orthodontistry [2,3], pain [4], cancer [5-7], and mental health [8,9], among many others. Such studies often look at information available to, and used by, people in particular geographic areas,

for example, pediatric asthma in Saudi Arabia [10], preconception care in Italy [11], and medical information in Brazil [12] and Portugal [13]. A PubMed search for research into online medical information, including, for example, use of resources such as Wikipedia or Google in medical education and availability of information for patients, provides thousands of search results. This is indicative of the fact that the Internet has become a source of medical information for patients and health care professionals alike, as shown by the increasing prevalence of the Internet use and social networking associated with “Web 2.0” for information sourcing and sharing online [14].

In the area of generic medicines, misconceptions and misinformation exist that are easily disseminated and perpetuated online. Given that health care professionals have expressed poor opinions of generics in the past [15], it is therefore challenging to communicate accurate information to the general public about the medicines that they are taking. There is a need to provide accurate information, to dispel myths, and to counter misinformation, but also to present the material in a manner that is accessible to the intended audience. For example, it has been reported that, in the case of patients particularly, myths and uncertainties about generic medicines abound and that accurate information can be difficult to come by [16].

A good quality medical or health care information website could be defined as one that contains accurate and unbiased information on all aspects of the topic (both positive and negative) for which the website is published, in conjunction with the ability of the website to be easily read and understood by its target audience. Where the audience is intended to be the general public, readability of the website will be a key factor in its success (as defined by the number of hits the website receives, indicative of its ranking on Internet search engine results). After all, if a website contains exemplary information but cannot be easily read and understood by its audience, it is possible for it to go largely undiscovered in the plethora of information available on the Internet. This study focused on non-advertised or promoted websites (ie, rankings on a Search Engine Results Page (SERP) that are not there as a result of a paid advertisement or promotion but rather are ranked and returned by Google’s algorithms).

While the availability and accuracy of existing online medical/health care information continues to be studied, much less work appears to have been performed in the area of development of medical information websites—in particular websites aimed at providing accurate and unbiased medical information to the general public. A PubMed search done February 22, 2013, using the search term *development medical information website*, returned 28 articles specifically related to the topic of development of medical/health care information websites.

The objective of this paper was to provide a method for the planning of information to be included in medical information websites and for representing that information in a readable

manner. As Search Engine Optimization (SEO) can be a critical factor in ensuring top-ranking search engine results [17] and given that the cost of using potentially expensive online advertising or SEO professionals in order to promote a website may be prohibitive for government or advocacy groups wishing to impart good quality medical/health care information, use of the tools and techniques described in this paper will not only ensure the quality of the information in the website but may also provide the website with an improved chance of being returned to a searcher in a higher ranking on a Google SERP, without incurring significant additional cost.

Methods

Rationale

To ensure a high-quality medical information website, two factors should be considered in its development: (1) the information it will present (quality, accuracy, comprehensiveness, balance, impartiality, etc) and (2) the ability of the information to be read and understood by the target audience.

Based on these factors, an assessment tool was developed that may be used to prospectively design the content of an optimized website. This study reports the composition of that tool and its validation through retrospective assessment of existing sites.

Information Gathering and Website Quality Assessment Tool Development

A tool for assessment of websites imparting information on generic drugs was developed. This Website Quality Assessment (WQA) tool consisted of a series of yes/no type questions, where a point was awarded for positive or correct information (see Table 1). No points were awarded for information lacking or for inaccurate information. Questions that cannot be answered were designated “not applicable” (N/A) and no score awarded. An overall WQA score for each website was totaled from the scores assigned to each assessment question.

In the development of the WQA tool, the following criteria were used:

1. Is there a listing of the questions likely to be asked by the searcher?
2. What myths or misinformation exist on the topic that may need to be dispelled or corrected?
3. What information could be required by the searcher in order to assist in making informed decisions?
4. Are there relevant comparisons or analogies that might help in understanding of the topic by a nonscientist or clinician?
5. Is there any associated or corollary information from other related topics or areas that might be helpful to support understanding of the topic?

The number of assessment questions will be determined by the topic in question and is not fixed. However, all areas in the 5 criteria steps noted above should be covered in the WQA questions used.

Table 1. Website Quality Assessment for assessing information on websites on generic medicines.

Question	Answer and score
Does the site explain what a generic medicine is?	Yes=1 No=0
Is this explanation correct?(ie, equivalent in dose, strength, route of administration, safety, efficacy, and intended use)	Yes=1 No=0
If so, is the explanation of a generic medicine readable and understandable by a nonscientist?	Yes=1 No=0
Are examples given of generic medicines? Eg, example of a proprietary medicine that also states the counterpart generic medicine?	Yes=1 No=0
Is bioequivalence mentioned in the website?	Yes=1 No=0
Is bioequivalence explained?	Yes=1 No=0 N/A
If so, is the explanation of bioequivalence correct?	Yes=1 No=0 N/A
If so, is the explanation of bioequivalence readable and understandable by a nonscientist?	Yes=1 No=0 N/A
Is the cheaper price of generics mentioned?	Yes=1 No=0
Is an accurate reason for the cheaper price of generics given?	Yes=1 No=0 N/A
Is any inaccurate information regarding the cheaper price of generics given?	Yes=0 No=1 N/A
Are examples given of the actual price difference between generics and proprietary medicines, or of the amount of money that can be saved by use of generics?	Yes=1 No=0
Is reference made to the fact that approved, equivalent generic meds can have a different appearance (color, shape, etc) different taste/smell or different inactive ingredients?	Yes=1 No=0
Are narrow therapeutic index (NTI) drugs mentioned?	Yes=1 No=0
Is the difference between NTI and non-NTI drugs explained?	Yes=1 No=0 N/A
Is there accurate information given on how generic bioequivalence or generic manufacturing may affect NTI drugs?	Yes=1 No=0 N/A
Is any inaccurate information given regarding NTI drugs?	Yes=0 No=1 N/A
Are “pros” of generics mentioned? (eg, lower price for same safety and bioequivalence, etc)	Yes=1 No=0
Are any “cons” of generics mentioned? (eg, adverse events to dissimilar excipients, etc)	Yes=1 No=0
Is the difference between proprietary and nonproprietary names mentioned?	Yes=1 No=0
Is the explanation given for the difference between proprietary and nonproprietary names accurate?	Yes=1 No=0 N/A
Is generic prescribing mentioned and explained accurately?	Yes=1 No=0
Total WQA score	
Flesch Reading Ease score	
Flesch-Kincaid Grade Level	

Validation of the WQA Tool

To validate the tool, all searches were performed on Google (google.com) and a number of the resulting hits in the SERPs returned were assessed using the 22-question Generic Medicines WQA (Table 1). The search was physically done in several English-speaking countries, using computers with Internet protocol (IP) addresses in those countries, in order to determine if there was any country-to-country (or geographic) variability. The searches were performed in the United States, Canada, Ireland, Great Britain, and Australia. The search term used was identical in all cases: “generic drug OR medicine” (without the quotes). All searches were performed during March and April of 2012, and a total of 24 distinct websites were assessed.

To measure reproducibility of use of the tool, each of the websites was independently assessed by 2 different raters.

Assessment of Website Readability

Readability of text is an important issue, especially in the medical domain. For this study readability of text was assessed using two methods: (1) Flesch Reading Ease score and (2) Flesch-Kincaid Grade Level. However, it is worth noting that other readability evaluation methods have also been used in the assessment of medical texts [18].

A minimum of a 100-word sample of continuous text was selected at random from the website text and pasted into Microsoft Word. This text was then analyzed using the readability statistics in the MS Word application.

MS Word’s Flesch Reading Ease score is based on a formula developed in 1948 by Rudolf Flesch [19]. It is computed using the average number of syllables per word and words per sentence. Syllables-per-word is a measure of word difficulty. Words-per-sentence is an indicator of syntactic complexity.

The Flesch Reading Ease scale ranges from zero to 100. Zero to 50 is very difficult to difficult reading. Eighty and above is easy to very easy reading. Flesch himself set the minimum score for plain English at 60 [19]. Microsoft's documentation encourages authors of standard documents to aim for a score of 60 to 70 [20,21].

The Flesch-Kincaid Grade Level, which was developed in 1975, measures the readability of a document based on the minimum education level required for a reader to understand it [22]. Microsoft recommends aiming for a Flesch-Kincaid score of 7.0 to 8.0 for most documents. According to a 1993 study, the average adult in the United States reads at the seventh-grade level and the authors of that study recommended that materials for the public be written at a fifth- or sixth-grade reading level [20].

Statistical Analyses

The mean and standard deviation of the differences between the 2 reviewers for all three tools (WQA, Flesch Reading Ease score, and Flesch Kincaid Grade Level) were used to calculate limits of agreement, which are represented graphically in Bland-Altman plots. The intraclass correlation coefficient (ICC) was used to measure reproducibility. Spearman correlation coefficient (r_s) was used to measure the association between the ranking of websites with WQA scores and readability assessments. Absolute values of $r_s > 0.3$ were considered to represent moderate correlations; > 0.5 were considered strong correlations. The scores from the developer of the assessment tool (SD) were used in the correlation analyses. The correlation between ranking of websites and WQA scores was also used to demonstrate the predictive validity of this newly developed assessment tool.

Results

Validation of the WQA Tool

Statistical analysis of the 2 independent raters (SSD and NC) using Bland-Altman plots showed that, for WQA assessments of the websites, the mean difference (SSD minus NC) represented by the solid black line in a) in [Multimedia Appendix 1](#) was zero (SD 1.18) indicating perfect agreement on average. The median difference was also zero (range -3 to 2). Only one observation was outside the limits of agreement (this website was a list of brand name medicines alongside the names of their generic counterparts). One rater performed the WQA based on this list, whereas the second rater looked for information on other pages of the website, thus accounting for the difference in WQA ratings awarded. An ICC value of 0.94 indicated excellent reproducibility between different users.

Similar analysis of the readability of the websites using Flesch Reading Ease score (on a scale of 0 to 100) and Flesch-Kincaid Grade Level (on a scale of 1 to 18) showed comparable levels of agreement (see b) and c) in [Multimedia Appendix 1](#). The mean difference (SSD minus NC) for reading ease score is 4.66

(SD 12.06) indicating that rater SSD was scoring slightly higher than NC on average. The mean difference (rater SSD minus NC) for grade level was -1.79 (SD 2.86) indicating that rater SSD was scoring slightly lower than NC on average. One observation in each case was outside the limits of agreement. However, as each rating was independent, different sections of text were likely to be taken from each of the websites assessed. This variation in the text taken most likely accounted for the single observation outside the limits of agreement. An ICC value of 0.71 for Flesch Reading Ease score and 0.63 for Flesch-Kincaid Grade Level demonstrate moderate to strong reproducibility, particularly given the subjectivity of this type of assessment, and the possible variability in the text selected by reviewers for assessment.

Overall, the WQA and readability scores demonstrate acceptable reproducibility of the tools when by used by more than 1 rater.

Correlation Between WQA Score and SERP Ranking

Scatterplots of WQA score against rankings on Google SERPs in different regions worldwide (United States, Canada, Ireland, United Kingdom, and Australia) are given in [Multimedia Appendix 2](#). Using Spearman correlation coefficient, a moderate to strong correlation between a WQA score and ranking on Google SERPs could be seen ([Table 2](#)). The observed relationship was seen in Google searches done in the different regions worldwide indicating that the correlation occurs regardless of the location or IP address of the searcher's computer. The strongest correlation ($r_s = -0.67$), was seen in the Google search performed in the United States.

Therefore, use of WQA assessment questions while developing information for inclusion in a medical information website could, by corollary, be a step towards ensuring higher Google SERP rankings and, therefore, exposure to a greater potential audience for the website.

Correlation of Readability With SERP Ranking

There was also a relationship, in general, between readability and ranking on Google searches ([Table 2](#)). Flesch Reading Ease scores were correlated with the SERP ranking of the websites in each country. Again, the strongest relationship was seen in the US Google search ($r_s = -0.64$). In general, the top ranked sites (placed 1, 2, etc) tended to have the higher Reading Ease scores. Because of the small sample sizes in the study (at most 10 websites in each domain) and hence low statistical power, a descriptive analysis is presented and no hypothesis tests were carried out.

Additionally, scores for Flesch-Kincaid Grade Level assessments were correlated with SERP ranking of the websites. In general, the top ranked sites tended to have lower grade level values with the most significant relationship again being seen in the US search (r_s value of 0.68). Therefore, the implication is that that websites with greater ease of readability are more likely to rank high in, and therefore be accessed from, Google SERPs.

Table 2. Correlation between WQA, reading ease score, and grade level with ranking using Spearman correlation coefficient (r_s).

Domain	n	WQA, Spearman r_s	Flesch Reading Ease score, Spearman r_s	Flesch-Kincaid Grade Level, Spearman r_s
US / .com	7	-0.67	-0.64	0.68
CA / .com	8	-0.38	-0.48	0.43
IE / .com	8	-0.49	-0.33	0.24
UK / .com	8	-0.38	-0.48	0.43
AU / .com	8	-0.34	0.29	-0.38

Discussion

Principal Findings

Prior to publication of a website, information must be gathered and written that will be disseminated to the intended audience through the website. Development and use of a specific WQA-type assessment during the design phase of a medical/health care information website on any topic will ensure that the information put into the website is of sufficient quality to satisfy potential searchers and users of the website. WQA can be used to assess drafts of the information to be published. Use of positive and negative scoring (positive scoring for information that is necessary, of good quality, and needed to support the integrity of the website; negative scoring for any information that is inaccurate, biased, or that may take from the integrity of the information) employed by WQA assessment ensures that all aspects of the information gathering initiative are accounted for during the website design.

As the Internet is one of the first places a patient is likely to go when searching for medical information [23] and given that Google is the primary search engine in use worldwide, holding almost 90% of the global search engine market [24], corollary use of WQA could possibly lead to higher rankings on Google SERPs for websites using this tool in their design and development.

Furthermore, this study has demonstrated that websites with greater ease of readability are more likely to rank high in, and therefore be accessed from, Google searches. Therefore, inclusion of Flesch Reading Ease and Flesch-Kincaid Grade Level assessments as part of the WQA enable a more comprehensive assessment of how the website might perform in Google searches. We have demonstrated in this paper that high readability scores and WQA scores are more likely to lead to a high Google SERP ranking.

Limitations

A limitation of this study is the small number of websites assessed. Further studies in this area could make use of technology, for example, a web crawler to gain additional information that could allow for clustering or commonalities across a spectrum of similar websites to be examined. A further study could evaluate sites containing similar content but focus instead on usability and accessibility, for example, are the sites well designed, are they pleasing to the eye, and is the navigation user-friendly? Isolating such content from the design and visual presentation of websites would provide further insight into the usability and accessibility of medical information providing

websites that would complement the findings in this paper. Indeed, information from such a study, if done using websites focused on generic medicines, may provide insight into the adoption and penetration of such medicines in different markets worldwide.

Readability formulas, additionally, have limitations in that a favorable score may not always be fully indicative of clarity of information (for instance, not all low-syllable words are always clearly understood, shorter sentences are not always necessarily easier to read, and inferences may be required that may increase the complexity of the text). Therefore, these formulas need to be used in conjunction with other plain language guidelines when writing for provision of health care information (especially for low literacy and limited English proficiency audiences), and not used as sole measures of understandability.

Conclusions

With about 16% of adults in the United Kingdom being described as “functionally literate” (ie, they have literacy levels at or below those expected from an 11-year old [25]), and the International Adult Literacy Survey showing that 1 in 4 adults in the Republic of Ireland have problems with even the simplest of literacy tasks [26] (with similar rates being seen in the United States [27] and Canada [28]), it is fair to say that writing of medical information websites with this in mind may be the most important aspect in providing medical information to the general public. This point, of course, applies to all printed material (eg, pamphlets given to patients), not just information published online. Arguably, it follows that training writers of medical information (to be disseminated to the general public, for instance) in methods of presenting simple, clear language is an important aspect in ensuring that the general public understand the information that health care professionals might be trying to impart to them. This becomes particularly important in light of research showing that there is often a discrepancy between the information that a physician believes a patient to have and what the patient actually understands [29].

Language complexity as a block to accessibility of information has been recognized by Wikipedia, the 6th most commonly accessed website in the world [30] and, as a solution, Wikipedia is available in both English and Simple English, where the Simple version is intended to be more accessible by use of simplified language and limited vocabulary. Consequently, Wikipedia guidelines on writing of the Simple version may be of use to those creating medical information websites for the general public [31].

Overall, use of the WQA tool in the planning and preparation of material for medical information websites, alongside an assessment of readability of the written material, is likely to ensure that the website subsequently ranks higher in Google SERPs and is thus more likely to be accessed, as well as read and understood, by the intended audience.

Acknowledgments

The authors would like to thank Ms YT Chueh and Dr Phil Hensche for their help in performing the Internet searches. This work was supported in part by a scholarship from the Faculty of Education and Health Sciences, University of Limerick, Ireland.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bland-Altman plots for WQA, Flesch Reading Ease Score, and Flesch Kincaid grade level.

[[PDF File \(Adobe PDF File\), 257KB - jmir_v15i8e183_app1.pdf](#)]

Multimedia Appendix 2

Scatterplots of WQA score against rankings on .com domains.

[[PDF File \(Adobe PDF File\), 260KB - jmir_v15i8e183_app2.pdf](#)]

References

1. Langille M, Bernard A, Rodgers C, Hughes S, Leddin D, van Zanten SV. Systematic review of the quality of patient information on the internet regarding inflammatory bowel disease treatments. *Clin Gastroenterol Hepatol* 2010 Apr;8(4):322-328. [doi: [10.1016/j.cgh.2009.12.024](#)] [Medline: [20060070](#)]
2. Livas C, Delli K, Ren Y. Quality evaluation of the available Internet information regarding pain during orthodontic treatment. *Angle Orthod* 2013 May;83(3):500-506. [doi: [10.2319/052512-435.1](#)] [Medline: [23025687](#)]
3. Fraval A, Ming Chong Y, Holcdorf D, Plunkett V, Tran P. Internet use by orthopaedic outpatients - current trends and practices. *Australas Med J* 2012;5(12):633-638 [FREE Full text] [doi: [10.4066/AMJ.2012.1530](#)] [Medline: [23382767](#)]
4. Colón Y. Searching for pain information, education, and support on the Internet. *J Pain Palliat Care Pharmacother* 2013 Mar;27(1):71-73. [doi: [10.3109/15360288.2012.758218](#)] [Medline: [23394472](#)]
5. Peterson MW, Fretz PC. Patient use of the internet for information in a lung cancer clinic. *Chest* 2003 Feb;123(2):452-457. [Medline: [12576365](#)]
6. Helft PR, Eckles RE, Johnson-Calley CS, Daugherty CK. Use of the internet to obtain cancer information among cancer patients at an urban county hospital. *J Clin Oncol* 2005 Aug 1;23(22):4954-4962. [doi: [10.1200/JCO.2005.09.621](#)] [Medline: [16051948](#)]
7. Vordermark D, Kölbl O, Flentje M. The Internet as a source of medical information. Investigation in a mixed cohort of radiotherapy patients. *Strahlenther Onkol* 2000 Nov;176(11):532-535. [Medline: [11143529](#)]
8. Powell J, Clarke A. Internet information-seeking in mental health: population survey. *Br J Psychiatry* 2006 Sep;189:273-277 [FREE Full text] [doi: [10.1192/bjp.bp.105.017319](#)] [Medline: [16946364](#)]
9. Lissman TL, Boehnlein JK. A critical review of internet information about depression. *Psychiatr Serv* 2001 Aug;52(8):1046-1050. [Medline: [11474049](#)]
10. AlSaadi MM. Evaluation of internet use for health information by parents of asthmatic children attending pediatric clinics in Riyadh, Saudi Arabia. *Ann Saudi Med* 2012;32(6):630-636. [doi: [10.5144/0256-4947.2012.630](#)] [Medline: [23396028](#)]
11. Agricola E, Gesualdo F, Pandolfi E, Gonfiantini MV, Carloni E, Mastroiacovo P, et al. Does googling for preconception care result in information consistent with international guidelines: a comparison of information found by Italian women of childbearing age and health professionals. *BMC Med Inform Decis Mak* 2013;13:14 [FREE Full text] [doi: [10.1186/1472-6947-13-14](#)] [Medline: [23347453](#)]
12. Gondim AP, Weyne DP, Ferreira BS. Quality of health and medication information on Brazilian websites. *Einstein (Sao Paulo)* 2012 Sep;10(3):335-341 [FREE Full text] [Medline: [23386014](#)]
13. Del Giglio A, Abdala B, Ogawa C, Amado D, Carter D, Gomieiro F, et al. Quality of internet information available to patients on websites in Portuguese. *Rev Assoc Med Bras* 2012;58(6):645-649 [FREE Full text] [Medline: [23250091](#)]
14. 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](#)] [Medline: [23153164](#)]

15. Dunne S, Shannon B, Dunne C, Cullen W. A review of the differences and similarities between generic drugs and their originator counterparts, including economic benefits associated with usage of generic medicines, using Ireland as a case study. *BMC Pharmacol Toxicol* 2013;14:1 [FREE Full text] [doi: [10.1186/2050-6511-14-1](https://doi.org/10.1186/2050-6511-14-1)] [Medline: [23289757](https://pubmed.ncbi.nlm.nih.gov/23289757/)]
16. Baumgärtel C. Myths, questions, facts about generic drugs in the EU. *GaBI J* 2012 Feb 15;1(1):34-38. [doi: [10.5639/gabij.2012.0101.009](https://doi.org/10.5639/gabij.2012.0101.009)]
17. Auinger A, Brandtner P, Großdeßner P, Holzinger A. Findability and Usability as Key Success Factors. In: Proceedings of the International Conference on e-Business (DCNET/ICE-B/OPTICS). 2012 Presented at: International Conference on e-Business (DCNET/ICE-B/OPTICS); July 24-27, 2012; Rome.
18. Holzinger A, Baerenthaler M, Pammer W, Katz H, Bjelic-Radisic V, Ziefle M. Investigating paper vs. screen in real-life hospital workflows: Performance contradicts perceived superiority of paper in the user experience. *International Journal of Human-Computer Studies* 2011 Aug;69(9):563-570. [doi: [10.1016/j.ijhcs.2011.05.002](https://doi.org/10.1016/j.ijhcs.2011.05.002)]
19. Flesch R. A new readability yardstick. *J Appl Psychol* 1948 Jun;32(3):221-233. [Medline: [18867058](https://pubmed.ncbi.nlm.nih.gov/18867058/)]
20. Stockmeyer NO. Using Microsoft Word's Readability Program. 2009. URL: <http://www.michbar.org/journal/pdf/pdf4article1467.pdf> [accessed 2013-03-17] [WebCite Cache ID 6FBcYBYCa]
21. Test your document's readability. URL: <http://office.microsoft.com/en-us/word-help/test-your-document-s-readability-HP010148506.aspx> [accessed 2013-03-17] [WebCite Cache ID 6FBcwjzei]
22. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Springfield, Virginia: National Technical Information Service; 1975. URL: <http://digitalcollections.lib.ucf.edu/cdm4/document.php?CISOROOT=/IST&CISOPTR=26301&CISOSHOW=26253> [WebCite Cache ID 6FBe9CuPV]
23. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients' use of the Internet for medical information. *J Gen Intern Med* 2002 Mar;17(3):180-185 [FREE Full text] [Medline: [11929503](https://pubmed.ncbi.nlm.nih.gov/11929503/)]
24. StatCounter. Bing Overtakes Yahoo! Globally for First Time.: *BusinessWire*; 2011. URL: <http://www.businesswire.com/news/home/20110301006261/en/Bing-Overtakes-Yahoo!-Globally-Time-%D0-StatCounter> [accessed 2013-03-17] [WebCite Cache ID 6FBd5fjZo]
25. How many illiterate adults are there in England?. URL: http://www.literacytrust.org.uk/adult_literacy/illiterate_adults_in_england [accessed 2013-03-17] [WebCite Cache ID 6FBdF9w4m]
26. Literacy in Ireland. URL: <http://www.nala.ie/literacy/literacy-in-ireland> [accessed 2013-03-17] [WebCite Cache ID 6FBdJIGwy]
27. IALS Results. National Center for Education Statistics URL: <http://nces.ed.gov/surveys/all/results.asp> [accessed 2013-03-17] [WebCite Cache ID 6FBdNgMBd]
28. Adult Literacy in OECD Countries: Technical Report on the First International Adult Literacy Survey. URL: <http://nces.ed.gov/pubs98/98053.pdf> [accessed 2013-03-17] [WebCite Cache ID 6FBdTjIDu]
29. Olson DP, Windish DM. Communication discrepancies between physicians and hospitalized patients. *Arch Intern Med* 2010 Aug 9;170(15):1302-1307. [doi: [10.1001/archinternmed.2010.239](https://doi.org/10.1001/archinternmed.2010.239)] [Medline: [20696951](https://pubmed.ncbi.nlm.nih.gov/20696951/)]
30. Top 500 Global Sites. 2013. URL: <http://www.alexa.com/topsites> [accessed 2013-03-17] [WebCite Cache ID 6FBdafLXI]
31. Wikipedia: How to write Simple English pages. URL: http://simple.wikipedia.org/wiki/Wikipedia:How_to_write_Simple_English_pages [accessed 2013-03-17] [WebCite Cache ID 6FBdd9Tea]

Abbreviations

ICC: intraclass correlation coefficient

IP: Internet protocol

SEO: search engine optimization

SERP: search engine results page

WQA: website quality assessment

Edited by G Eysenbach; submitted 25.03.13; peer-reviewed by A Holzinger, F Bassetti; comments to author 12.05.13; revised version received 17.05.13; accepted 11.06.13; published 27.08.13.

Please cite as:

Dunne S, Cummins NM, Hannigan A, Shannon B, Dunne C, Cullen W

A Method for the Design and Development of Medical or Health Care Information Websites to Optimize Search Engine Results Page Rankings on Google

J Med Internet Res 2013;15(8):e183

URL: <http://www.jmir.org/2013/8/e183/>

doi: [10.2196/jmir.2632](https://doi.org/10.2196/jmir.2632)

PMID: [23981848](https://pubmed.ncbi.nlm.nih.gov/23981848/)

©Suzanne Dunne, Niamh Maria Cummins, Ailish Hannigan, Bill Shannon, Colum Dunne, Walter Cullen. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 27.08.2013. 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

An Analysis of Online Evaluations on a Physician Rating Website: Evidence From a German Public Reporting Instrument

Martin Emmert¹, MSc, PhD; Florian Meier¹, MSc

Institute of Management (IFM), School of Business and Economics, Friedrich-Alexander-University Erlangen-Nuremberg, Nuremberg, Germany

Corresponding Author:

Martin Emmert, MSc, PhD

Institute of Management (IFM)

School of Business and Economics

Friedrich-Alexander-University Erlangen-Nuremberg

Lange Gasse 20

Nuremberg, 90403

Germany

Phone: 49 911 5302 ext 253

Fax: 49 911 5302 114

Email: Martin.Emmert@fau.de

Abstract

Background: Physician rating websites (PRW) have been gaining in popularity among patients who are seeking a physician. However, little evidence is available on the number, distribution, or trend of evaluations on PRWs. Furthermore, there is no published evidence available that analyzes the characteristics of the patients who provide ratings on PRWs.

Objective: The objective of the study was to analyze all physician evaluations that were posted on the German PRW, jameda, in 2012.

Methods: Data from the German PRW, jameda, from 2012 were analyzed and contained 127,192 ratings of 53,585 physicians from 107,148 patients. Information included medical specialty and gender of the physician, age, gender, and health insurance status of the patient, as well as the results of the physician ratings. Statistical analysis was carried out using the median test and Kendall Tau-b test.

Results: Thirty-seven percent of all German physicians were rated on jameda in 2012. Nearly half of those physicians were rated once, and less than 2% were rated more than ten times (mean number of ratings 2.37, SD 3.17). About one third of all rated physicians were female. Rating patients were mostly female (60%), between 30-50 years (51%) and covered by Statutory Health Insurance (83%). A mean of 1.19 evaluations per patient could be calculated (SD 0.778). Most of the rated medical specialties were orthopedists, dermatologists, and gynecologists. Two thirds of all ratings could be assigned to the best category, "very good". Female physicians had significantly better ratings than did their male colleagues ($P<.001$). Additionally, significant rating differences existed between medical specialties ($P<.001$). It could further be shown that older patients gave better ratings than did their younger counterparts ($P<.001$). The same was true for patients covered by private health insurance; they gave more favorable evaluations than did patients covered by statutory health insurance ($P<.001$). No significant rating differences could be detected between female and male patients ($P=.505$). The likelihood of a good rating was shown to increase with a rising number of both physician and patient ratings.

Conclusions: Our findings are mostly in line with those published for PRWs from the United States. It could be shown that most of the ratings were positive, and differences existed regarding sociodemographic characteristics of both physicians and patients. An increase in the usage of PRWs might contribute to reducing the lack of publicly available information on physician quality. However, it remains unclear whether PRWs have the potential to reflect the quality of care offered by individual health care providers. Further research should assess in more detail the motivation of patients who rate their physicians online.

(*J Med Internet Res* 2013;15(8):e157) doi:[10.2196/jmir.2655](https://doi.org/10.2196/jmir.2655)

KEYWORDS

physician rating website; public reporting; quality of care; Internet; patient satisfaction

Introduction

In many health care systems, quality of care improvement strategies have been implemented over the last few years [1]; nevertheless, quality deficits still remain [2-4]. Several studies have further shown remarkable variability in quality of care across health care providers [1,5-7]. However, patients are not likely to be generally aware of existing quality differences [8,9]. One reason for this is the limited amount of publicly reported information on the quality of health care providers [10].

It has become a major challenge to remedy this deficiency by improving transparency about the quality of health care providers [10,11]. This is supposed to increase overall quality by steering patients to better performing health care providers [12,13] and by motivating providers to make quality improvements [9,14]. Therefore, public reporting (PR) instruments have been put in place in many countries [15-22]. These instruments generally assess the quality of care by measuring adherence to clinical guidelines and by providing additional structural information [11]. However, patients have been slow to take advantage of these comparative reports in making their health care provider choices [9]. Possible reasons for this might be found in the fact that patients are not aware of the information, do not understand it, do not believe it, or are unwilling or unable to use the information provided [23].

The newest trend in the PR movement is the use of physician rating websites (PRWs) [24]. The primary objective of these websites lies in rating and discussing physician quality online by using user-generated data [25,26]. Although the usefulness of PRWs has been seen critically from a scientific point of view [24], their popularity among patients has been increasing [24,27,28]. In contrast to traditional PR instruments, PRWs might have the advantage that the information can be more easily understood by patients. While traditional instruments report on measures such as the administration of beta blockers or angiotensin-converting enzyme inhibitors, which require a higher level of clinical knowledge than most patients have [8], PRWs concentrate on measuring patient satisfaction [24].

Although there is a vast amount of evidence regarding traditional PR instruments, little research has addressed PRWs [25]. A recently conducted systematic review has identified 9 articles published in peer-reviewed journals [25]. In them, the number, distribution, and trend of the evaluations on PRWs were investigated [11,27-34]. Most of the investigations evaluated ratings for a (non)random sample of physicians, while 1 study assessed over 386,000 national ratings from 2005 to 2010 from the US PRW, RateMDs. Furthermore, there is no published evidence available that analyzes the characteristics of the patients who provide ratings.

In this context, this paper adds to the literature by presenting an analysis of all physician evaluations posted on the German PRW, jameda, in 2012. Thereby, we provide descriptive analysis of (1) both physician and patient characteristics, and (2) the number, distribution, and results of the ratings. Analytical analyses were applied to assess (3) the impact of physician and patient characteristics on the overall performance measure, and

(4) the correlation between the number of ratings per patient/physician and the overall performance.

Methods

Analysis of Jameda

This paper presents an analysis of all 127,192 physician evaluations that were posted on the German PRW, jameda, in 2012. In total, 107,148 patients completed evaluations on 53,585 physicians. The dataset contained the following information: the medical specialty and gender of the physician, as well as the gender, age, and health insurance status of the patient. Additionally, the results of the physician ratings for all mandatory and optional questions were included. The mandatory physician rating system on jameda consists of 5 questions, rated according to the grading system in German schools on a 1-6 scale (1=very good; 2=good; 3=satisfactory; 4=fair; 5=deficient; and 6=insufficient) [35]. These relate to (Q1) satisfaction with the treatment offered by the physician, (Q2) education about the illness and treatment, (Q3) the relationship of trust with the physician, (Q4) the time the physician spent on the patient's concerns, and (Q5) the friendliness of the physician. A mean score ("overall performance") is calculated, based on the results of these 5 questions. Beyond that, a narrative commentary has to be given and 13 optional questions are available for answering (these are not addressed in this paper) [36].

We focused on jameda because it is likely to play the most significant role in the German PRW movement for the following reasons: (1) from a patient's perspective, jameda is the PRW to which a patient is most likely to be referred [24,31], (2) jameda is ranked highest in traffic among German PRWs [34], and (3) among German PRWs, jameda has been shown to contain the largest number of ratings, so far [37].

Statistical Analysis

All statistical analyses were conducted using SPSS 21.0 (SPSS for Windows, version 21.0). The median test was used for nonparametric data of groups with different distributions. The Kendall Tau-b test was used to analyze specific correlations. Differences were considered to be significant if $P < .05$ and highly significant if $P < .001$.

Results

Number and Distribution of Ratings

In total, 127,192 ratings of 53,585 physicians from 107,148 patients were posted on the PRW, jameda, in 2012. The German outpatient sector consists of approximately 146,000 physicians [38]; thus, 37% were rated in 2012. As displayed in Table 1, about one third of all rated physicians were female (34.1%). The rating patients were mostly female (60%), between 30-50 years (51%), and covered by Statutory Health Insurance (83%).

The distribution of ratings demonstrates that nearly half of the physicians were rated once and less than 2% were rated more than ten times (see Table 2). Thereby, rated physicians had a mean of 2.37 individual ratings (SD 3.169, range 1-159). It could further be shown that 88% of the patients left a single rating and 12% of them left between two and five ratings. This

leads to an average of 1.19 rated physicians per patient (SD 0.778, range 1-153).

If the ratings are analyzed according to the medical specialty of the physicians in absolute terms, family physician/general practitioner, internist, and gynecologist were rated most often (13,466, 8709, and 6410, respectively) (see Table 3; [38,39]). In contrast, laboratory specialist, nuclear medicine, and child

and youth psychotherapist were rated least frequently (13, 136, and 166, respectively). The distribution of ratings in relative terms, compared to the national physician composition, shows that the most rated medical specialties were orthopedists, dermatologists, and gynecologists (59.20%, 58.90%, and 56.90%, respectively). In contrast, the least frequently rated medical specialties were radiologists, anesthetists, and laboratory specialists (10.40%, 7.90%, and 2.10%, respectively).

Table 1. Number and distribution of ratings on jameda (gender, age, insurance).

Characteristics	Absolute	%	%, cum
Gender—Physician			
Female	18,284	34.1	34.1
Male	35,301	65.9	100.0
Total	53,585	100.0	
Gender—Patient			
Female	48,171	45.0	45.0
Male	31,809	29.7	74.7
n.a.	27,168	25.4	100.0
Total	107,148	100.0	
Age—Patient			
<30	13,639	12.7	12.7
30-50	38,608	36.0	48.8
50+	23,676	22.1	70.9
n.a.	31,225	29.1	100.0
Total	107,148	100.0	
Health insurance—Patient			
Statutory health insurance	64,986	60.7	60.7
Private health insurance	13,402	12.5	73.2
n.a.	28,760	26.8	100.0
Total	107,148	100.0	

Table 2. Number and distribution of ratings on jameda (physicians and patients).

	Number of ratings	Absolute	%	%, cum
Physicians				
	1	26,615	49.7	49.7
	2-5	23,430	43.7	93.4
	6-10	2,664	5.0	98.4
	11-50	849	1.6	99.9
	51+	27	0.1	100.0
	Total	53,585	100.0	
Patients				
	1	94,099	87.8	87.8
	2-5	12,702	11.9	99.7
	6-10	329	0.3	100.0
	11-51	17	0.0	100.0
	51+	1	0.0	100.0
	Total	107,148	100.0	

Table 3. Number and distribution of ratings according to medical specialty.

Medical specialty	Rated physicians in absolute terms (%)	Number of physicians in Germany ^a	Rated physicians in relative terms (%)
Orthopedist	3677 (6.9)	6206	59.2
Dermatologist (incl venereologist)	2445 (4.6)	4154	58.9
Gynecologist	6410 (12.0)	11,256	56.9
Oral maxillo-facial surgeon	634 (1.2)	1,122	56.5
Neurosurgeon	337 (0.6)	608	55.4
ENT specialist, otorhinolaryngologist	2304 (4.3)	4301	53.6
Urologist	1545 (2.9)	3030	51.0
Neurologist/Psychiatrist	2685 (5.0)	5775	46.5
Pediatrician	2957 (5.5)	6866	43.1
Medical practitioner without specialization	1697 (3.2)	4252	39.9
Ophthalmologist	2253 (4.2)	5796	38.9
Internist	8709 (16.3)	23,198	37.5
Family physician/General practitioner	13,466 (25.1)	36,196	37.2
Nuclear medicine	136 (0.3)	698	19.5
Child and youth psychotherapist	166 (0.3)	922	18.0
Others	3432 (6.4)	23,561 ^{b,c}	14.6
Radiologist (incl radiotherapist)	421 (0.8)	4,029	10.4
Anesthetist	298 (0.6)	3796	7.9
Laboratory specialist	13 (0.0)	623	2.1
Total	53,585 (100.0)	146,389	36.6

^aIf not other than [38].

^bAccording to [39].

^cOthers (eg, surgeon, psychotherapist, pathologist, pneumologist).

Evaluations

Table 4 shows the evaluation results of all 53,585 rated physicians (as they are displayed on the website). It can be shown that two thirds of all evaluations were assigned to the best rating category, “very good”. An additional 13% of patients rated their experience with the physician as “good”. Three percent of the physicians were rated with the worst score, “insufficient” in their overall performance. The median result of all questions was “very good”, while the mean varied between 1.68 for question 5 (friendliness of the physician) and 1.85 for question 3 (relationship of trust with the physician).

An analysis was performed to ascertain whether differences in the rating of a physician, regarding both the physician (ie, gender and medical specialty) and the patient characteristics (ie, gender, age, and health insurance) could be determined. The results are displayed in Table 5. They show that female physicians were rated better than their male colleagues and that the difference is statistically significant (the percentage of rated physicians below median is 61% for female and 59% for male physicians; $P<.001$). Furthermore, significant rating differences between medical specialties could be demonstrated ($P<.001$). The best rated medical specialties were laboratory specialists, anesthetists, medical practitioner without specialization, and family physician/general practitioner (85%, 76%, 74%, and 70% below median, respectively). The lowest ratings were given to neurologist/psychiatrist, ophthalmologist, orthopedist, and

dermatologist (including venereologist) (47%, 45%, 35%, and 35% below median, respectively).

With respect to patient characteristics, no significant rating differences between female and male patients could be detected (percentage below median is 59% in each group; $P=.505$). However, it could be shown that older patients gave better ratings than did their younger counterparts ($P<.001$). Additionally, patients covered by private health insurance gave more favorable evaluations than did patients covered by statutory health insurance ($P<.001$).

Next, the correlation between the mean overall performance of a physician and the number of ratings per physician was addressed. As displayed in Figure 1, the total performance range can be observed for physicians with a low number of ratings. By contrast, physicians who received a higher number of ratings were shown to have better ratings (eg, all physicians with more than 60 ratings were rated as “very good”). As a result, the correlation between the mean overall performance of a physician and the number of ratings per physician could be shown to be statistically significant (Kendall Tau-b=0.193, $P<.001$). This is also true for all five mandatory questions ($P<.001$; data not presented here). We further investigated to find out whether similar results could be detected for the number of ratings per patient compared to the mean overall performance given by this patient. The result is displayed in Figure 1 and shows a similar correlation (Kendall Tau-b=0.178, $P<.001$).

Table 4. Evaluation results of all rated physicians on jameda.

	Overall performance	Q1 ^b	Q2 ^c	Q3 ^d	Q4 ^e	Q5 ^f
Performance range^a, n (%)						
1	35,227 (65.7)	35,030 (65.4)	33,345 (62.2)	34,665 (64.7)	34,331 (64.1)	36,708 (68.5)
2	7170 (13.4)	7302 (13.6)	8660 (16.2)	6748 (12.6)	7535 (14.1)	7313 (13.6)
3	4694 (8.8)	4876 (9.1)	5019 (9.4)	5077 (9.5)	5075 (9.5)	4305 (8.0)
4	2615 (4.9)	2312 (4.3)	2584 (4.8)	2350 (4.4)	2512 (4.7)	2201 (4.1)
5	2429 (4.5)	2000 (3.7)	1988 (3.7)	1972 (3.7)	1992 (3.7)	1461 (2.7)
6	1450 (2.7)	2065 (3.9)	1988 (3.7)	2773 (5.2)	2139 (4.0)	1597 (3.0)
Total	53,585 (100.0)	53,585 (100.0)	53,584 (100.0)	53,585 (100.0)	53,584 (100.0)	53,585 (100.0)
Mean	1.77	1.79	1.83	1.85	1.82	1.68
Median	1.00	1.00	1.00	1.00	1.00	1.00
SD	1.32	1.35	1.34	1.43	1.37	1.24
Minimum	1.00	1.00	1.00	1.00	1.00	1.00
Maximum	6.00	6.00	6.00	6.00	6.00	6.00

^aGerman school based rating system (1=very good; 2=good; 3=satisfactory; 4=fair; 5=deficient; 6=insufficient).

^bQ1: satisfaction with the treatment by the physician.

^cQ2: education about the illness and treatment.

^dQ3: relationship of trust with the physician.

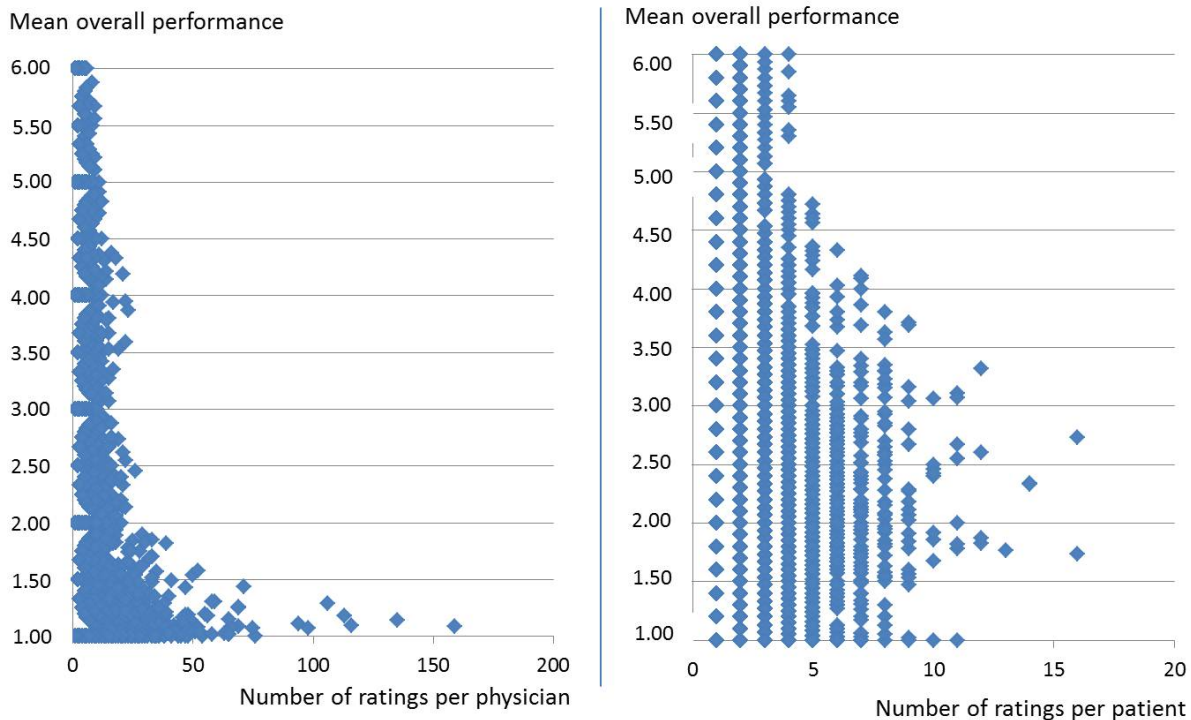
^eQ4: time the physician spent for the patient’s concerns.

^fQ5: friendliness of the physician.

Table 5. Ratings differences regarding physician and patient characteristics.

Characteristics	N	> Median	≤Median	Percentage below median (%)	P value ^a
Gender—Physician					<.001
Female	18,284	7168	11,116	61	
Male	35,301	14,583	20,718	59	
Medical specialty—Physician					<.001
Laboratory specialist	13	2	11	85	
Anesthetist	298	72	226	76	
Medical practitioner without specialization	1697	433	1264	74	
Family physician/General practitioner	13,466	4081	9385	70	
Oral maxillo-facial surgeon	634	197	437	69	
Internist	8709	2889	5820	67	
Nuclear medicine	136	47	89	65	
Urologist	1545	562	983	64	
Neurosurgeon	337	124	213	63	
Pediatrician	2957	1180	1777	60	
Radiologist (incl radiotherapist)	421	174	247	59	
Gynecologist	6410	2662	3748	58	
Others	3432	1504	1928	56	
Child and youth psychotherapist	166	78	88	53	
ENT specialist, otorhinolaryngologist	2304	1108	1196	52	
Neurologist/Psychiatrist	2685	1424	1261	47	
Ophthalmologist	2253	1241	1012	45	
Orthopedist	3677	2380	1297	35	
Dermatologist (incl venereologist)	2445	1593	852	35	
Gender—Patient					.505
Female	48,171	19,989	28,182	59	
Male	31,809	13,124	18,685	59	
Age—Patient					<.001
<30	13,639	6697	6942	51	
30-50	38,608	16,064	22,544	58	
50+	23,676	8542	15,134	64	
Health Insurance—Patient					<.001
Statutory Health Insurance	64,986	28,309	36,677	56	
Private Health Insurance	13,402	4523	8879	66	

^aMedian test.

Figure 1. Scatterplot (bivariate); the number of ratings per physician (left)/patient (right) with the mean overall performance for a rated physician.

[note that one outlier (N=153 ratings) was not integrated into this figure]

Discussion

Principal Findings

In this section, the results obtained in this investigation are compared to published studies, mostly from the United States. The evidence from this investigation shows that 37% of physicians in the German outpatient sector were rated on jameda in 2012. This number exceeded those from previously published international studies. For example, Gao and colleagues showed that 16% of US physicians received an online review on RateMDs in the period between 2005 and 2010 [27]. Lagu et al reported that out of 300 Boston physicians, 27% of them had been rated [11], while Mostaghimi et al calculated percentages of between 0.4% and 21% for a sample of 250 randomly selected internal medicine physicians [33]. In a sample of 500 randomly selected US urologists, the percentages varied between 0.4% and 53.6% [40]. Published results for German PRWs reported percentages of between 3.36% and 25.78% in 2009 [31] and between 3% and 28% in 2012 [34]. However, it is worth mentioning here that direct comparison is difficult due to the fact that data from one year was analyzed in this investigation, whereas most studies use ratings for a sample of physicians without including any time constraints.

It could also be shown that rated physicians had a mean of 2.37 individual ratings (SD 3.169, range 1-159). Published results for the US PRW, RateMDs, were quite similar and were reported to be 2.7 [30], respectively 3.2 [27]. More recent US studies determined numbers of 2.35 [11] and 2.4 [40], while results for German PRWs were reported to be between 1.1 and 3.9 [34]. The number decreases to 0.87 when regarding all rated physicians from the German outpatient sector in 2012. This is

slightly higher than the results obtained by Lagu and colleagues (mean 0.63) [11].

Nearly half of the physicians were rated only once, and 44% received between 2 and 5 ratings in this study. Less than 2% were rated more than 10 times and 0.1% more than 50 times. These numbers are in line with the results obtained by analyzing the ratings provided for 2010 on RateMDs. In that case, half of the physicians had a single rating and the percentage of physicians with 5 or more ratings was 12.50% [27]. Of 250 randomly selected physicians in Boston, 50 physicians (20%) had between 1 and 4 reviews on Healthgrades, 13 physicians (5.2%) on RateMDs, and 1 physician (0.4%) on Wellness. Only 3 physicians had more than 5 reviews on any of the ratings sites [33].

About one third of all rated physicians on jameda were female. This is consistent with both the gender composition of physicians in Germany (female national average 40% [38]) and with the results by Gao and colleagues [27]. If the ratings are analyzed according to the medical specialty in relative terms (ie, compared to the national physician composition), the numbers are again confirmed by other study results. For example, Gao and colleagues showed that rated physicians were most likely to be classified as obstetrician/gynecologists and least likely to be classified as other specialists such as radiologists or anesthesiologists [27].

In this study, almost 80% of all evaluations could be assigned to the two best rating categories. Less than 3% of the physicians were rated with the worst score, "insufficient". These results are in line with most other studies: Lagu and colleagues categorized 88% of quantitative reviews as positive, 6% as negative, and 6% as neutral [11]. On RateMDs, 45.80% of the

physicians received the best score and only 12% were rated with the worst score [27]. Kadry et al assessed the 10 most commonly visited US PRWs and found that the percentage of reviews rated ≥ 75 on a 100-point scale was 61.5%, ≥ 4 on a 5-point scale was 57.74%, and ≥ 3 on a 4-point scale was 74.0% [32]. On the Canadian PRW RateMDs, 70% of the comments were reported to be favorable and about 30% of the comments were negative [41]. In the sample of 500 randomly selected US urologists, 86% had positive ratings [40]. Moreover, the median result of all questions in this study was "very good". The means varied between 1.68 concerning the friendliness of the physician (question 5) and 1.85 regarding the relationship of trust with the physician (question 3). In their study, Kadry et al determined the average rating to be 77 out of 100 for sites using a 100-point scale, 3.84 out of 5 for sites using a 5-point scale, and 3.1 out of 4 for sites using a 4-point scale [32]. For the US RateMDs, the mean scores were reported to be 3.93 [27] and 3.82 [30] on a 5-point scale, respectively. Finally, a comprehensive analysis of German PRWs showed the mean ratings to be between 1.1 and 1.5 (3-point scale, 1 "good", 3 "poor") [34].

The results of this study suggest that female physicians receive better ratings than do their male colleagues. The number is small but statistically significant ($P < .001$). Better ratings for female physicians were also determined by Ellimoottil and colleagues ($P = .72$) [40]. However, this is in contrast to the results obtained by Gao and colleagues, who showed that male physicians received higher ratings than did female physicians ($P < .001$) [27]. But, differences in all three studies were shown to be quite small.

We can further demonstrate significant rating differences among the analyzed medical specialties. Of these, the best rated were laboratory specialists, anesthesiologists, medical practitioners without specialization, and family physician/general practitioners. The lowest ratings were given to neurologist/psychiatrists, ophthalmologists, orthopedists, and dermatologists. In line with the numbers obtained in this study, higher ratings were shown for physicians in primary care [27] and lower ratings for physicians in dermatology [30]. However, in another study, primary care physicians were rated at average [30]. Lagu et al found a similar percentage of positive, negative, and neutral quantitative reviews for generalists and subspecialists. They then concluded that after accounting for varying number of reviews per physician, generalists tended to have more positive reviews than did subspecialists [11].

This is the first study that allows for a closer analysis of the patients who rate their physicians. Approximately 73% of all patients provided information regarding gender, age, and health insurance. According to our results, most of the rating patients were female (60%) and were covered by Statutory Health Insurance (83%). One other notable fact could be shown: patients in the youngest age group (< 30) made fewer ratings than did older patients. Whether or not this is due to more severe illness problems with increasing age cannot be assessed with this data. However, this question should be addressed in future research.

The fact that hardly any patients leave more than a single rating (mean 1.19 rated) can be regarded as even more surprising. One

might expect that once they were aware of the existence of such websites, patients would use them constantly in an active (ie, rating physicians) or passive (ie, only searching for physicians) manner, especially to assist other patients with information when seeking a physician. However, we could not investigate the motivation behind the patients' ratings. Nor could we assess the reasons for not regularly rating physicians. Considering the mean of 14 [42] to 17 [43] physician contacts in Germans with statutory health insurance, there is still high potential for even more ratings. The fact that patients covered by private health insurance give more favorable ratings than do patients covered by statutory health insurance is not surprising, since they were found to have faster access to care [44]. This might well have had an effect on the ratings differences. Whether quality of care differences can be determined between the two groups and whether this leads to ratings differences should be addressed in future studies.

It could be shown that there is a significant correlation between the mean overall performance rating of a physician and the number of ratings received for that physician ($P < .001$). One possible explanation for this finding might be the fact that physicians who are aware of these websites and use them as a marketing instrument may specifically ask satisfied patients to leave a (positive) rating on a PRW. Another explanation might be that some physicians, who are identified by patients on PRWs, simply provide outstanding quality of care and they receive favorable ratings afterwards. Although our results prove that there is a significant correlation between these variables, we cannot prove which assumption is true. This should be addressed in further studies, which should contain additional information about the physicians.

Limitations

There are some limitations that have to be taken into account when interpreting the results of this investigation. First, we analyzed online ratings from only a single PRW, jameda. Although jameda has shown to be the most frequently used German PRW, it is possible that other PRWs have more online reviews or show other results. Second, the data provided allowed for comprehensive analysis. However, there was no information available on the age of the physician, malpractice claims, or the medical school attended. This information would have allowed further analysis. Third, we were not able to present analysis conducted over a longer period of time. However, the data do reflect the entire year 2012. Fourth, we did not analyze results presented in narrative comments. Finally, there was no chance to verify the validity of the analyzed reviews. Therefore, it cannot be guaranteed that the ratings were not subject to manipulation [27].

Conclusions

Finally, it can be stated that there is a limited amount of publicly reported information on quality of health care providers. To increase transparency, different approaches have been developed. There are traditional PR instruments that focus on the adherence to evidence-based guidelines. Thus, they may have the potential to reflect the clinical quality of care provided by a health care professional. However, these instruments have not yet proven to be a meaningful measure for patients. In

contrast, PRWs concentrate on patient satisfaction measures. Whether or not these results have the potential to reflect the quality of care provided by a health care professional should be addressed in future research as well. Since an increasing usage of these websites has already been shown [24,27,28], PRWs might contribute to reducing the lack of publicly available

information on quality, at least for those physicians who have been rated. Given that only a certain number of physicians has been rated so far, there is still no perfect transparency. However, given the increasing number of ratings on PRWs, the future impact for patients seeking a physician will continue to rise.

Acknowledgments

The authors would like to thank Dr. Phillip Goos and Ms Elke Ruppert from jameda for the provision of the data. The abstract was accepted for publication at the Medicine 2.0'13 congress (Sept 23-24, 2013).

Conflicts of Interest

None declared.

References

1. Lebrun LA, Shi L, Zhu J, Sharma R, Sripipatana A, Hayashi AS, et al. Racial/ethnic differences in clinical quality performance among health centers. *J Ambul Care Manage* 2013;36(1):24-34. [doi: [10.1097/JAC.0b013e3182473523](https://doi.org/10.1097/JAC.0b013e3182473523)] [Medline: [23222010](https://pubmed.ncbi.nlm.nih.gov/23222010/)]
2. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003 Jun 26;348(26):2635-2645. [doi: [10.1056/NEJMsa022615](https://doi.org/10.1056/NEJMsa022615)] [Medline: [12826639](https://pubmed.ncbi.nlm.nih.gov/12826639/)]
3. Asch SM, Kerr EA, Keesey J, Adams JL, Setodji CM, Malik S, et al. Who is at greatest risk for receiving poor-quality health care? *N Engl J Med* 2006 Mar 16;354(11):1147-1156. [doi: [10.1056/NEJMsa044464](https://doi.org/10.1056/NEJMsa044464)] [Medline: [16540615](https://pubmed.ncbi.nlm.nih.gov/16540615/)]
4. Laschet H. *ÄrzteZeitung*. 2013 Jan 21. GBA knöpft sich fünf Problem-Indikation vor URL: http://www.aerztezeitung.de/politik_gesellschaft/versorgungsforschung/article/829958/qualitaet-gba-knoepft-fuenf-problem-indikation.html [accessed 2013-07-17]
5. Dynan L, Goudie A, Smith RB, Fairbrother G, Simpson LA. Differences in quality of care among non-safety-net, safety-net, and children's hospitals. *Pediatrics* 2013 Feb;131(2):304-311. [doi: [10.1542/peds.2012-1089](https://doi.org/10.1542/peds.2012-1089)] [Medline: [23296439](https://pubmed.ncbi.nlm.nih.gov/23296439/)]
6. Merchant RM, Yang L, Becker LB, Berg RA, Nadkarni V, Nichol G, American Heart Association Get With the Guideline-Resuscitation Investigators. Variability in case-mix adjusted in-hospital cardiac arrest rates. *Med Care* 2012 Feb;50(2):124-130 [FREE Full text] [doi: [10.1097/MLR.0b013e31822d5d17](https://doi.org/10.1097/MLR.0b013e31822d5d17)] [Medline: [22249921](https://pubmed.ncbi.nlm.nih.gov/22249921/)]
7. Tsai C, Sullivan AF, Gordon JA, Kaushal R, Magid DJ, Blumenthal D, et al. Racial/ethnic differences in emergency care for joint dislocation in 53 US EDs. *Am J Emerg Med* 2012 Nov;30(9):1970-1980. [doi: [10.1016/j.ajem.2012.04.023](https://doi.org/10.1016/j.ajem.2012.04.023)] [Medline: [22795991](https://pubmed.ncbi.nlm.nih.gov/22795991/)]
8. Hibbard J, Sofaer S. Best Practices in Public Reporting No 1: How To Effectively Present Health Care Performance Data To Consumers. AHRQ Publications No. 10-0082-EF 2010:-.
9. Hibbard JH. What can we say about the impact of public reporting? Inconsistent execution yields variable results. *Ann Intern Med* 2008 Jan 15;148(2):160-161. [Medline: [18195340](https://pubmed.ncbi.nlm.nih.gov/18195340/)]
10. Porter ME, Guth C. Chancen für das deutsche Gesundheitssystem. Berlin: Springer/Gabler; 2012.
11. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. *J Gen Intern Med* 2010 Sep;25(9):942-946 [FREE Full text] [doi: [10.1007/s11606-010-1383-0](https://doi.org/10.1007/s11606-010-1383-0)] [Medline: [20464523](https://pubmed.ncbi.nlm.nih.gov/20464523/)]
12. Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality-of-care information? A systematic review. *Med Care* 2009 Jan;47(1):1-8. [doi: [10.1097/MLR.0b013e3181808bb5](https://doi.org/10.1097/MLR.0b013e3181808bb5)] [Medline: [19106724](https://pubmed.ncbi.nlm.nih.gov/19106724/)]
13. Emmert M, Gemza R, Schöffski O, Sohn S. [Public reporting in health care: the impact of publicly reported quality data on patient steerage]. *Gesundheitswesen* 2012 Jun;74(6):e25-e41. [doi: [10.1055/s-0031-1285857](https://doi.org/10.1055/s-0031-1285857)] [Medline: [21866496](https://pubmed.ncbi.nlm.nih.gov/21866496/)]
14. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 2008 Jan 15;148(2):111-123. [Medline: [18195336](https://pubmed.ncbi.nlm.nih.gov/18195336/)]
15. Mukamel DB, Mushlin AI. Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State Cardiac Surgery Mortality Reports. *Med Care* 1998 Jul;36(7):945-954. [Medline: [9674613](https://pubmed.ncbi.nlm.nih.gov/9674613/)]
16. Chassin MR. Achieving and sustaining improved quality: lessons from New York State and cardiac surgery. *Health Aff (Millwood)* 2002;21(4):40-51 [FREE Full text] [Medline: [12117152](https://pubmed.ncbi.nlm.nih.gov/12117152/)]
17. Mukamel DB, Weimer DL, Zwanziger J, Gorthy SF, Mushlin AI. Quality report cards, selection of cardiac surgeons, and racial disparities: a study of the publication of the New York State Cardiac Surgery Reports. *Inquiry* 2004;41(4):435-446. [Medline: [15835601](https://pubmed.ncbi.nlm.nih.gov/15835601/)]
18. Stevenson DG. Is a public reporting approach appropriate for nursing home care? *J Health Polit Policy Law* 2006 Aug;31(4):773-810. [doi: [10.1215/03616878-2006-003](https://doi.org/10.1215/03616878-2006-003)] [Medline: [16971545](https://pubmed.ncbi.nlm.nih.gov/16971545/)]

19. NHS. Quality and Outcomes Framework. 2013. Online GP practice results database: Search for the QOF results for your GP practice URL: <http://www.qof.ic.nhs.uk/search/index.asp> [accessed 2013-07-17] [WebCite Cache ID 61B8B1MxE]
20. NHS Choices. Your health, your choices. 2013. URL: <http://www.nhs.uk/Pages/HomePage.aspx>; [accessed 2013-07-17] [WebCite Cache ID 61B8DGgP]
21. Wubker A, Sauerland D, Wubker A. Beeinflussen bessere Qualitätsinformationen die Krankenhauswahl in Deutschland? Eine empirische Untersuchung (Does Better Quality Information Affect Hospital Choice in Germany? An Empirical Analysis). *Jahrbuecher fur Nationaloekonomie und Statistik* 2010;230:467-490.
22. Geraedts M, Schwartze D, Molzahn T. Hospital quality reports in Germany: patient and physician opinion of the reported quality indicators. *BMC Health Serv Res* 2007;7:157 [FREE Full text] [doi: [10.1186/1472-6963-7-157](https://doi.org/10.1186/1472-6963-7-157)] [Medline: [17903244](https://pubmed.ncbi.nlm.nih.gov/17903244/)]
23. Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. *Annu Rev Public Health* 2003;24:413-433. [doi: [10.1146/annurev.publhealth.24.100901.141005](https://doi.org/10.1146/annurev.publhealth.24.100901.141005)] [Medline: [12428034](https://pubmed.ncbi.nlm.nih.gov/12428034/)]
24. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. *Methods Inf Med* 2012;51(2):112-120. [doi: [10.3414/ME11-01-0045](https://doi.org/10.3414/ME11-01-0045)] [Medline: [22101427](https://pubmed.ncbi.nlm.nih.gov/22101427/)]
25. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. *J Med Internet Res* 2013;15(2):e24 [FREE Full text] [doi: [10.2196/jmir.2360](https://doi.org/10.2196/jmir.2360)] [Medline: [23372115](https://pubmed.ncbi.nlm.nih.gov/23372115/)]
26. Hardey M. Consuming Professions: User-review websites and health services. *Journal of Consumer Culture* 2010;10(1):129-149.
27. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012;14(1):e38 [FREE Full text] [doi: [10.2196/jmir.2003](https://doi.org/10.2196/jmir.2003)] [Medline: [22366336](https://pubmed.ncbi.nlm.nih.gov/22366336/)]
28. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. *J Gen Intern Med* 2012 Jun;27(6):685-692. [doi: [10.1007/s11606-011-1958-4](https://doi.org/10.1007/s11606-011-1958-4)] [Medline: [22215270](https://pubmed.ncbi.nlm.nih.gov/22215270/)]
29. Alemi F, Torii M, Clementz L, Aron DC. Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. *Qual Manag Health Care* 2012;21(1):9-19. [doi: [10.1097/QMH.0b013e3182417fc4](https://doi.org/10.1097/QMH.0b013e3182417fc4)] [Medline: [22207014](https://pubmed.ncbi.nlm.nih.gov/22207014/)]
30. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. *Inform Prim Care* 2009;17(4):249-253. [Medline: [20359403](https://pubmed.ncbi.nlm.nih.gov/20359403/)]
31. Emmert M, Maryschok M, Eisenreich S, Schöffski O. [Websites to assess quality of care--appropriate to identify good physicians?]. *Gesundheitswesen* 2009 Apr;71(4):e18-e27. [doi: [10.1055/s-0028-1103288](https://doi.org/10.1055/s-0028-1103288)] [Medline: [19221986](https://pubmed.ncbi.nlm.nih.gov/19221986/)]
32. Kadry B, Chu LF, Kadry B, Gammass D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. *J Med Internet Res* 2011;13(4):e95 [FREE Full text] [doi: [10.2196/jmir.1960](https://doi.org/10.2196/jmir.1960)] [Medline: [22088924](https://pubmed.ncbi.nlm.nih.gov/22088924/)]
33. Mostaghimi A, Crotty BH, Landon BE. The availability and nature of physician information on the internet. *J Gen Intern Med* 2010 Nov;25(11):1152-1156 [FREE Full text] [doi: [10.1007/s11606-010-1425-7](https://doi.org/10.1007/s11606-010-1425-7)] [Medline: [20544300](https://pubmed.ncbi.nlm.nih.gov/20544300/)]
34. Strech D, Reimann S. Deutschsprachige Arztbewertungsportale: Der Status quo ihrer Bewertungskriterien, Bewertungstendenzen und Nutzung: German Language Physician Rating Sites: The Status Quo of Evaluation Criteria, Evaluation Tendencies and Utilization. *Gesundheitswesen* 2012;74:502-503.
35. Geraedts M, Hermeling P, de Cruppé W. Communicating quality of care information to physicians: a study of eight presentation formats. *Patient Educ Couns* 2012 Jun;87(3):375-382. [doi: [10.1016/j.pec.2011.11.005](https://doi.org/10.1016/j.pec.2011.11.005)] [Medline: [22177585](https://pubmed.ncbi.nlm.nih.gov/22177585/)]
36. Reimann S, Strech D. The representation of patient experience and satisfaction in physician rating sites. A criteria-based analysis of English- and German-language sites. *BMC Health Serv Res* 2010;10:332 [FREE Full text] [doi: [10.1186/1472-6963-10-332](https://doi.org/10.1186/1472-6963-10-332)] [Medline: [21138579](https://pubmed.ncbi.nlm.nih.gov/21138579/)]
37. Emmert M, Sander U, Maryschok M, Esslinger AS, Schöffski O. Arzt-Bewertungsportale im Internet: Eine aktuelle Bestandsaufnahme. *IMPLICATIONplus - Gesundheitspolitische Analysen* 2010;9:-.
38. Bundesärztekammer. Ergebnisse der Ärztestatistik zum 31. 2011 Dec. URL: <http://www.bundesaerztekammer.de/page.asp?his=0.3.10275>; [accessed 2013-07-17] [WebCite Cache ID 61B8zOc1R]
39. Kassenärztliche Bundesvereinigung, National Association of Statutory Health Insurance Physicians. Statistische Informationen aus dem Bundesarztregister: Bundesgebiet insgesamt. 2012. URL: <http://daris.kbv.de/daris.asp>
40. Ellimoottil C, Hart A, Greco K, Quek ML, Farooq A. Online reviews of 500 urologists. *J Urol* 2013 Jun;189(6):2269-2273. [doi: [10.1016/j.juro.2012.12.013](https://doi.org/10.1016/j.juro.2012.12.013)] [Medline: [23228385](https://pubmed.ncbi.nlm.nih.gov/23228385/)]
41. Mackay B. RateMDs.com nets ire of Canadian physicians. *CMAJ* 2007 Mar 13;176(6):754 [FREE Full text] [doi: [10.1503/cmaj.070239](https://doi.org/10.1503/cmaj.070239)] [Medline: [17353522](https://pubmed.ncbi.nlm.nih.gov/17353522/)]
42. Grobe T, Bitzer E, Schwartz F. Barmer GEK Arztreport 2013: Schwerpunkt. Hannover: ADHS; 2013.
43. Riens B, Erhard M, Mangiapane S. Primary care physician contacts in 2007 & #8211; Background and analyses. Berlin; 2012. URL: [http://www.versorgungsatlas.de/fileadmin/ziva_docs/14/Arztkontakte_Abstract%20\(englisch\).pdf](http://www.versorgungsatlas.de/fileadmin/ziva_docs/14/Arztkontakte_Abstract%20(englisch).pdf) [accessed 2013-07-17] [WebCite Cache ID 61B9r7G9M]
44. Roll K, Stargardt T, Schreyögg J. Effect of Type of Insurance and Income on Waiting Time for Outpatient Care. Working Papers, Hamburg Center for Health Economics, University of Hamburg 2011:-.

Abbreviations**PR:** public reporting**PRW:** physician rating website

Edited by G Eysenbach; submitted 07.04.13; peer-reviewed by B Kadry, E Santoro, Q Zhang; accepted 11.06.13; published 06.08.13.

Please cite as:

Emmert M, Meier F

An Analysis of Online Evaluations on a Physician Rating Website: Evidence From a German Public Reporting Instrument

J Med Internet Res 2013;15(8):e157

URL: <http://www.jmir.org/2013/8/e157/>

doi: [10.2196/jmir.2655](https://doi.org/10.2196/jmir.2655)

PMID: [23919987](https://pubmed.ncbi.nlm.nih.gov/23919987/)

©Martin Emmert, Florian Meier. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 06.08.2013. 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

Physician Choice Making and Characteristics Associated With Using Physician-Rating Websites: Cross-Sectional Study

Martin Emmert¹, MSc, PhD; Florian Meier¹, MSc; Frank Pisch¹, BSc; Uwe Sander², MD

¹Institute of Management (IFM), School of Business and Economics, Friedrich-Alexander-University Erlangen-Nuremberg, Nuremberg, Germany

²University of Applied Sciences and Arts, Hannover, Germany

Corresponding Author:

Martin Emmert, MSc, PhD

Institute of Management (IFM)

School of Business and Economics

Friedrich-Alexander-University Erlangen-Nuremberg

Lange Gasse 20

Nuremberg, 902403

Germany

Phone: 49 9115302 ext 253

Fax: 49 911 5302 14

Email: Martin.Emmert@wiso.uni-erlangen.de

Abstract

Background: Over the past decade, physician-rating websites have been gaining attention in scientific literature and in the media. However, little knowledge is available about the awareness and the impact of using such sites on health care professionals. It also remains unclear what key predictors are associated with the knowledge and the use of physician-rating websites.

Objective: To estimate the current level of awareness and use of physician-rating websites in Germany and to determine their impact on physician choice making and the key predictors which are associated with the knowledge and the use of physician-rating websites.

Methods: This study was designed as a cross-sectional survey. An online panel was consulted in January 2013. A questionnaire was developed containing 28 questions; a pretest was carried out to assess the comprehension of the questionnaire. Several sociodemographic (eg, age, gender, health insurance status, Internet use) and 2 health-related independent variables (ie, health status and health care utilization) were included. Data were analyzed using descriptive statistics, chi-square tests, and *t* tests. Binary multivariate logistic regression models were performed for elaborating the characteristics of physician-rating website users. Results from the logistic regression are presented for both the observed and weighted sample.

Results: In total, 1505 respondents (mean age 43.73 years, SD 14.39; 857/1505, 57.25% female) completed our survey. Of all respondents, 32.09% (483/1505) heard of physician-rating websites and 25.32% (381/1505) already had used a website when searching for a physician. Furthermore, 11.03% (166/1505) had already posted a rating on a physician-rating website. Approximately 65.35% (249/381) consulted a particular physician based on the ratings shown on the websites; in contrast, 52.23% (199/381) had not consulted a particular physician because of the publicly reported ratings. Significantly higher likelihoods for being aware of the websites could be demonstrated for female participants ($P<.001$), those who were widowed ($P=.01$), covered by statutory health insurance ($P=.02$), and with higher health care utilization ($P<.001$). Health care utilization was significantly associated with all dependent variables in our multivariate logistic regression models ($P<.001$). Furthermore, significantly higher scores could be shown for health insurance status in the unweighted and Internet use in the weighted models.

Conclusions: Neither health policy makers nor physicians should underestimate the influence of physician-rating websites. They already play an important role in providing information to help patients decide on an appropriate physician. Assuming there will be a rising level of public awareness, the influence of their use will increase well into the future. Future studies should assess the impact of physician-rating websites under experimental conditions and investigate whether physician-rating websites have the potential to reflect the quality of care offered by health care providers.

(*J Med Internet Res* 2013;15(8):e187) doi:[10.2196/jmir.2702](https://doi.org/10.2196/jmir.2702)

KEYWORDS

physician-rating website; public reporting; patient satisfaction; physician choice making

Introduction

Several studies have demonstrated variability in the quality of care across health care providers (eg, [1-5]). However, because of the limited amount of publicly reported quality information [6], patients are not likely to be aware of such differences [7,8]. To overcome this situation, public reporting instruments have been put in place within the past few years (eg, [9-16]). These instruments generally assess the quality of care by measuring adherence to clinical guidelines and provide some additional structural information [17].

Public reporting is supposed to increase the overall standard given by health care providers because it demonstrates which physicians use higher quality standards. This information steers patients to better performing providers; hence, motivating physicians to improve their overall quality outcomes [18,19]. In this context, previous systematic research has shown that public reporting has the potential to stimulate quality improvement outcomes at the hospital level. However, the effect on physicians or physician groups remains unaddressed [19]. Another review summarized the impact of 12 different public reporting instruments and included evidence from 21 studies, mostly from the United States. This study demonstrated that public reporting can be effective in directing patients when seeking a health care provider, especially for elective procedures [20]. Nevertheless, many authors state that patients have been slow to take advantage of comparative reports when making a health care provider choice (eg, [8]). Possible reasons for this are that patients are not aware of the information, do not understand it, do not believe it, or are not willing or able to use the information provided [21-23].

The newest trend in public reporting is physician-rating websites [24,25]. The primary objective of these relies on rating and discussing the physician's standards by using user-generated data [25,26]. Such sites have been established in many countries worldwide, such as the United States, England, Germany, Canada, Australia, New Zealand, and the Netherlands [14,17,24,27-33]. Recent research in this field has focused on the number, distribution, and trend of evaluations on physician-rating websites [17,24,27-34]. It could be shown that approximately 1 in 6 physicians has been rated so far, and that approximately 90% of all ratings were positive [25]. Based on this information, it is assumed that the use of physician-rating websites will increase [24,25].

Thus far, no peer-reviewed research has focused on the influence of physician-rating websites when choosing a physician in the outpatient sector [28]. It still remains uncertain whether these websites have an influence on patients seeking and selecting a physician. Furthermore, it remains unclear what key predictors are associated with the knowledge and the use of physician-rating websites. In this context, this paper adds to the literature by investigating the influence of German physician-rating websites on patients choosing a physician in

the outpatient sector and identifying the main predictors associated with the awareness and use of such sites.

Methods

This study was designed as a cross-sectional survey. An online panel (Tomorrow Focus Media Opinion Pool) was consulted within Germany in January 2013. The online panel consisted of 3052 respondents who received €1 per finished survey. The panel consisted of online users who agreed to receive survey invitations about society or media-related topics once per month. They obtain bonus points that can be used for online shopping (eg, Amazon, Zalando, Douglas) or donations. The online survey was provided within the Tomorrow Focus Media Opinion Pool network. Several online channels were used to recruit participants for the survey. All panel members were invited via email and newsletter to participate (the invitation contained a link to enter the online survey). Additionally, online banner advertising was applied within the Tomorrow Focus Media network-related websites.

A questionnaire was developed containing 28 questions, addressing topics related to physician-rating websites (see [Multimedia Appendix 1](#)). The questionnaire was piloted by 50 individuals to ensure the comprehensibility of the wording and internal validity; final adjustments were made accordingly. The questionnaire first asked for the participants' sources when seeking a physician. Participants were then asked 5 questions associated with the awareness and use of physician-rating websites, which served as our dependent yes/no variables and are described in the following (questions 1-3 included a list of the 9 leading German physician-rating websites for selection):

1. Have you ever heard of any of the following physician-rating websites? (awareness)
2. Have you ever searched for a physician on any of the following physician-rating websites? (searching)
3. Have you ever posted a rating on any of the following physician-rating websites? (rating)
4. Have you ever selected a particular physician based upon the publicly published results on any physician-rating website? (positive impact)
5. Have you ever not selected a particular physician based upon the publicly published results on any physician-rating website? (negative impact)

Further questions related to the importance of physician information provided on physician-rating websites, such as age, gender, medical devices, and number of patients treated by using a 5-point Likert scale from 1 (no importance at all) to 5 (very important). Following the specific physician-rating website questions, participants were asked a series of background sociodemographic questions (eg, age, gender, marital status, Internet use, and education). The survey ended with 2 health-related questions concerning the awareness of physician-rating websites and their usage. Health care utilization was measured in terms of the number of physician encounters within the past 6 months.

In addition to descriptive statistics, we used bivariate analyses (chi-square tests and *t* tests) to examine whether differences existed between those participants who stated that they were aware of or have used physician-rating websites and those who did not. Binary multivariate logistic regression was performed to identify the main predictors associated with the awareness and use of physician-rating websites (see previous description for the 5 dependent yes/no variables). Therefore, demographic data was dichotomized to contain group sizes to at least 20 participants in each subgroup [35]. For example, Internet use of the subgroups (ie, more than once a week, once a week, less than once a day, and once a month) was grouped as less than once a day. To ensure representativeness, the study sample was weighted for age, gender, and marital status according to the most recent data from the German Federal Statistical Office from 2011 [36]. Results from the logistic regression are presented for both the observed and the weighted sample.

Health status was measured by applying the World Health Organization (WHO) 5-item Well-being Index (WHO-5). The latter is comprised of 5 items, each rated on a 6-point Likert scale from 0 (not present) to 5 (constantly present); a raw score was calculated afterwards by summarizing the single scores. Higher scores reflect higher well-being status; conversely, poor well-being status is represented by a raw score below 13 or if the patient answered 0 to 1 on any of the 5 items [37].

All statistical analyses were conducted by using SPSS ver 21.0 (IBM Corp, Armonk, NY, USA). Observed differences were considered statistically significant if $P < .05$.

Results

A total of 1505 respondents completed online interviews (response rate 49.28%) averaging 11.7 minutes. Regarding the study sample, 857/1505 respondents were female (57.25%), most were covered by statutory health insurance (SHI; 1173/1505, 80.67%), and the overall mean age was 43.73 (SD 14.39) years (see Table 1 for an overview of the study population). In all, 316 respondents (32.63%) had more than 5 encounters with a health care provider within the 6-month period before the survey.

The following are the results of our 5 main dependent variables (see Table 2 for an overview of the results). Approximately one-third (483/1505, 32.09%) of all respondents were aware of German physician-rating websites. Regarding the relative distribution of age, the highest awareness percentage was for the age range 31 to 40 years (37.01%). The lowest awareness was seen in the youngest age group, younger than 20 years (15.87%). However, differences between age groups were not proven to be statistically significant ($P = .08$). This is also true for differences regarding education, employment, Internet use, and health status. Statistically significant higher awareness levels were shown for female respondents (35.71% vs 27.50%, $P < .001$), those who were widowed ($P = .012$), covered by SHI ($P = .02$), and those with higher health care utilization ($P < .001$).

In all, 25.32% (381/1505) of the respondents reported to have actively searched for a physician using a German physician-rating website. Once more, statistically significant higher percentages could be shown for female respondents ($P = .02$), those who were widowed ($P < .001$), covered by SHI ($P = .03$), and those with higher health care utilization ($P < .001$). The highest active search percentage was calculated for the age ranges 31 to 40 years (29.18%) and 61 to 70 years (28.89%), respectively. With respect to employment, higher percentages were calculated for those who were unemployed compared to their employed counterparts (31.5%, $P = .009$).

Every ninth interviewee (166/1505, 11.03%) had already posted a rating on a physician-rating website. In other words, every third respondent who was aware of physician-rating websites (166/483, 34.37%) had already rated a physician. Differences with respect to marital status ($P = .04$), health insurance coverage ($P = .04$), and health care utilization ($P < .001$) were statistically significant. No meaningful differences were calculated for age, gender, Internet use, or other characteristics.

According to our results, physician-rating websites seem to have a meaningful influence on choosing a physician. For those respondents who had sought a physician online (381/1505, 25.32%), 327 respondents made their decision for a particular physician based on ratings shown on the websites. Considering this represents only one-quarter of respondents, not everyone has performed an online search using physician-rating websites. A physician search can also be performed on search engines, which are likely to transfer the seeker to the results presented on specific physician-rating websites. Consequently, it has to be distinguished between those respondents who use physician-rating websites to search for physicians and those who do not. Specifically, 249 respondents (249/381, 65.35%) claimed to have performed an online search on a physician-rating website and their decision was influenced on the provided ratings. Furthermore, those of younger age groups (≤ 40 years) have been influenced positively by the publicly reported data ($P = .002$); the highest percentages were reported for the age groups 21 to 30 years (33.01%) and 31 to 40 years (24.56%), respectively.

Conversely, physician-rating websites can have a meaningful negative influence on a patient's choice. In our sample, 258 respondents (17.14%) did not consult a particular physician because of evaluation results on the websites. As mentioned previously, one has to distinguish between those respondents using physician-rating websites to search for physicians and those who do not. It was shown that 199 respondents (199/381, 52.23%) had performed an online search using a physician-rating website and made a subsequent decision against a particular physician. According to our results, younger study participants were significantly more influenced than their older counterparts were ($P < .001$). This was also true for female respondents (19.14% vs 14.69%, $P = .02$), those with higher education ($P < .001$), those employed ($P = .04$), and those with higher health care utilization ($P = .003$).

Table 1. Overview of study sample (N=1505).

Characteristics	Study sample
Age (years), mean (SD)	43.73 (14.39)
Age range (years), n (%)	
≤20	63 (4.50)
21-30	206 (14.70)
31-40	281 (20.06)
41-50	331 (23.63)
51-60	306 (21.84)
61-70	180 (12.85)
>70	34 (2.43)
Gender, n (%)	
Female	857 (57.25)
Male	640 (42.75)
Marital status, n (%)	
Married	713 (48.90)
Single	560 (38.41)
Divorced	149 (10.22)
Widowed	39 (2.47)
Education, n (%)	
High school	683 (46.62)
Technical university entrance qualification	196 (13.38)
Intermediate secondary school	345 (23.55)
Polytechnic secondary school	71 (4.85)
Secondary general school	148 (10.10)
Without school qualification	2 (0.14)
Others	20 (1.37)
Employment, n (%)	
Self-employed	145 (9.85)
Civil servants	68 (4.62)
Employee	720 (48.91)
Apprentices	24 (1.63)
Unemployed	54 (3.67)
Pensioners	202 (13.72)
High school students	68 (4.62)
Students (university/technical university)	92 (6.25)
Others	99 (6.73)
Health insurance, n (%)	
Statutory health insurance	1173 (80.67)
Private health insurance	275 (18.91)
No health insurance	6 (0.41)
Health care utilization,^a n (%)	
No treatment	138 (9.45)
1	245 (16.77)

Characteristics	Study sample
2	312 (21.36)
3	267 (18.28)
4	183 (12.53)
≥5	316 (21.63)
Internet use, n (%)	
>once a day	1252 (83.19)
once a day	178 (11.83)
>once a week	68 (4.52)
once a week	5 (0.33)
>once a day	1 (0.07)
once a month	1 (0.07)
WHO-5 health status^b	
Overall, mean (SD)	14.53 (5.55)
Poor (<13), n (%)	653 (44.30)
Good (>), n (%)	821 (55.70)

^aNumber of encounters within the past 6 months.

^bWHO-5 Well-being Index.

Table 2. Overview of survey results.

Variable	Awareness (n=483, 32.09%)		Searching (n=381, 25.32%)		Rating (n=166, 11.03%)		Positive impact (n=327, 21.73%)		Negative impact (n=258, 17.14%)	
	Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)	
		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a
Age (years), mean (SD)	44.17 (14.06)	.44	44.61 (13.81)	.18	45.15 (14.37)	.19	41.57 (13.90)	.003	41.80 (13.86)	.02
Age range (years), n (%)										
≤20	15.87	.08	4.76	.007	1.59	.14	22.22	.002	12.70	<.001
21-30	33.50		27.18		11.65		33.01		29.61	
31-40	37.01		29.18		12.81		24.56		16.73	
41-50	30.51		24.77		9.37		20.24		16.31	
51-60	32.03		24.51		11.44		19.28		14.71	
61-70	32.78		28.89		14.44		13.89		12.22	
>70	35.29		26.47		11.76		11.76		8.82	
Gender, n (%)										
Female	35.7	<.001	27.8	.02	10.5	.40	24.2	.01	19.1	.02
Male	27.5		22.3		11.9		18.8		14.7	
Marital status, n (%)										
Married	31.4	.01	25.4	<.001	11.4	.04	20.2	.37	15.3	.25
Single	30.7		22.0		8.8		22.5		18.6	
Divorced	36.2		33.6		15.4		26.2		20.1	
Widowed	55.6		52.8		19.4		25.0		22.2	
Education, n (%)										
High School	35.1	.27	28.0	.17	11.3	.29	23.9	.31	20.9	<.001
Technical university entrance qualification	28.1		22.4		12.2		20.4		15.8	
Intermediate secondary school	32.5		27.8		11.3		22.3		17.1	
Polytechnic secondary school	31.0		21.1		7.0		22.5		8.5	
Secondary general school	29.7		19.6		12.8		16.2		9.5	
Without school qualification	50.0		50.0		50.0		50.0		0.0	
Others	15.0		15.0		0.0		10.0		0.0	
Employment, n (%)										
Self-employed	35.2	.06	29.0	.009	9.0	.14	22.1	.02	19.3	.04
Civil servants	33.8		26.5		14.7		25.0		20.6	
Employee	32.4		26.7		11.4		24.3		19.6	
Apprentices	33.3		29.2		12.5		37.5		29.2	
Unemployed	42.6		31.5		11.1		25.9		16.7	
Pensioners	31.2		27.2		14.9		15.3		11.9	
High School Students	13.2		4.4		2.9		13.2		10.3	
Students (university/technical university)	32.6		23.9		5.4		20.7		15.2	
Others	34.3		20.2		10.1		14.1		10.1	

Variable	Awareness (n=483, 32.09%)		Searching (n=381, 25.32%)		Rating (n=166, 11.03%)		Positive impact (n=327, 21.73%)		Negative impact (n=258, 17.14%)	
	Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)		Mean (SD)/ n (%)	
		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a		<i>P</i> ^a
Health insurance										
Statutory health insurance	34.4	.02	27.4	.03	12.0	.04	23.2	.07	17.7	.79
Private health insurance	25.8		19.6		6.9		17.1		16.0	
No health insurance	16.7		16.7		0.0		33.3		16.7	
Health care utilization										
No treatment	20.3	<.001	13.8	<.001	3.6	<.001	10.1	<.001	8.7	.003
1	24.5		19.2		5.7		16.7		15.1	
2	34.0		26.9		12.2		20.5		14.4	
3	30.3		24.3		9.4		22.8		19.1	
4	34.4		27.9		14.2		25.1		19.7	
5+	43.4		34.8		17.4		30.4		22.8	
Internet use										
> once a day	33.2	.32	26.4	.33	11.4	.67	22.7	.33	17.9	.14
once a day	25.3		18.5		7.9		15.2		10.7	
> once a week	29.4		25.0		13.2		22.1		22.1	
once a week	40.0		20.0		0.0		20.0		0.0	
> once a day	0.00		0.0		0.0		0.0		0.0	
once a month	0.00		0.0		0.0		0.0		0.0	
WHO-5 health status										
Overall, mean (SD)	14.53 (5.59)	.97	14.43 (5.72)	.66	15.12 (6.24)	.15	14.55 (5.62)	.95	14.56 (5.36)	.94
Poor (<13), n (%)	31.4	.58	25.4	.95	10.3	.31	21.6	.71	16.8	.73
Good (>), n (%)	32.8		25.6		11.9		22.4		17.5	

^a*P* value was calculated using chi-square test or *t* test.

In connection with the demographic- and health-related variables, health care utilization was significantly associated with all dependent variables in our binary multivariate logistic regression models (see [Tables 3](#) and [4](#)). Additionally, gender, health insurance status, and health care utilization were all strongly associated with awareness of physician-rating websites. Awareness results were significantly higher in female (OR 0.75, 95% CI 0.57-0.98, *P*=.04), those insured by SHI (OR 0.63, 95% CI 0.42-0.94, *P*=.03), and those with a higher number of physician encounters (OR 4.16, 95% CI 2.34-7.38, *P*<.001). The awareness tended to be higher in widowed respondents, those with a higher education level, self-employed, frequently use the Internet, and those with a good health status. However, these differences were not statistically significant. It could further be shown that health insurance status and health care utilization were the only 2 independent variables which were proven to be strongly associated with the rating activity on physician-rating websites. Scores were significantly higher in participants insured by SHI (OR 0.48, 95% CI 0.25-0.92, *P*=.04) and those with a higher number of physician encounters (OR 7.47, 95% CI 2.21-25.27, *P*<.001). With respect to the last

dependent variables of interest (ie, being positively or negatively influenced in choosing a physician by the results on physician-rating websites), only health care utilization could be shown to be strongly associated.

[Tables 5](#) and [6](#) show the results of the weighted binary multivariate logistic regression models. After controlling for age, gender, and marital status (according to the German population in 2011 [[36](#)]), health care utilization and Internet use were shown to be significantly associated in all 5 models. Both education and health insurance status could further be shown to be strongly associated with searching for physicians on physician-rating websites (*P*<.05). As shown, a higher education level and being insured by SHI (OR 0.56, 95% CI 0.34-0.90, *P*<.05) indicate higher ratios. Furthermore, labor and health status were strongly associated with the rating activity on physician-rating websites. The same is true for marital status, which negatively influences the choice made when using the results on physician-rating websites. Only age and gender did not reach statistical significance in any of the weighted multivariate models.

Table 3. Independent factors associated with physician-rating website relevant issues (unweighted sample).

Variables	Awareness			Searching			Rating			Positive impact			Negative impact		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Age			.80			.87			.99			.99			.99
Gender															
Female	1.00			1.00			1.00			1.00			1.00		
Male	0.75	0.57, 0.98	.04	0.81	0.61, 1.08	.13	1.35	0.91, 1.99	.22	0.84	0.62, 1.14	.22	0.79	0.57, 1.11	.13
Marital status															
Married	1.00			1.00			1.00			1.00			1.00		
Single	1.02	0.73, 1.41	.74	0.85	0.60, 1.21	.47	0.88	0.54, 1.44	.99	1.01	0.70, 1.45	.87	1.13	0.76, 1.68	.42
Divorced	1.09	0.70, 1.68	.63	1.29	0.82, 2.01	.24	1.28	0.71, 2.30	.29	1.38	0.85, 2.22	.23	1.46	0.85, 2.51	.13
Widowed	2.24	0.97, 5.17	.08	2.92	1.26, 6.75	.02	1.70	0.57, 5.03	.23	1.32	0.50, 3.48	.82	2.58	0.97, 6.89	.07
Education															
Matura examination	1.00			1.00			1.00			1.00			1.00		
Technical university entrance qualification	0.74	0.49, 1.10	.16	0.80	0.52, 1.22	.30	1.01	0.58, 1.78	.90	0.94	0.60, 1.46	.78	0.83	0.51, 1.34	.53
Intermediate secondary school	0.83	0.59, 1.15	.26	0.90	0.64, 1.28	.57	0.86	0.53, 1.41	.52	0.93	0.64, 1.34	.83	0.79	0.53, 1.19	.25
Polytechnic secondary school	0.76	0.41, 1.42	.42	0.55	0.27, 1.09	.10	0.28	0.08, 0.99	.045	1.14	0.58, 2.24	.57	0.33	0.12, 0.89	.03
Secondary general school	0.76	0.47, 1.22	.25	0.58	0.34, 0.99	.04	1.15	0.61, 2.19	.92	0.70	0.39, 1.25	.30	0.44	0.22, 0.88	.03
Others	0.21	0.05, 0.98	.045	0.33	0.07, 1.54	.15	0.28	0.03, 2.82	.22	0.57	0.14, 2.24	.31	0.00	0.00,	.99
Labor															
Self-employed	1.00			1.00			1.00			1.00			1.00		
Civil servants	0.90	0.43, 1.87	.78	0.83	0.38, 1.80	.62	2.76	0.89, 8.58	.15	1.08	0.48, 2.42	.94	1.13	0.47, 2.71	.94
Employee	0.63	0.39, 1.01	.07	0.63	0.38, 1.04	.09	1.46	0.64, 3.36	.35	0.81	0.47, 1.38	.40	0.88	0.49, 1.57	.43
Apprentices	0.78	0.36, 1.69	.60	0.77	0.34, 1.73	.55	0.60	0.16, 2.25	.46	0.64	0.28, 1.48	.30	0.57	0.23, 1.41	.13
Unemployed	0.84	0.39, 1.83	.71	0.82	0.36, 1.85	.63	1.09	0.31, 3.90	.99	0.87	0.37, 2.08	.56	0.87	0.32, 2.33	.53
Pensioners	0.55	0.30, 1.01	.06	0.64	0.34, 1.20	.17	1.90	0.73, 4.92	.21	0.52	0.26, 1.07	.06	0.49	0.22, 1.08	.05
Others	0.71	0.37, 1.35	.43	0.47	0.23, 0.96	.047	1.46	0.50, 4.24	.37	0.44	0.20, 0.97	.05	0.41	0.17, 1.02	.07
Health insurance															
Statutory health insurance	1.00			1.00			1.00			1.00			1.00		
Private health insurance	0.63	0.42, 0.94	.03	0.61	0.40, 0.93	.04	0.48	0.25, 0.92	.04	0.70	0.44, 1.09	.14	0.81	0.50, 1.32	.41
Health care utilization															
No treatment	1.00		<.001	1.00		<.001	1.00		<.001	1.00		<.001	1.00		.003
1	1.80	0.99, 3.27	.05	1.71	0.89, 3.31	.10	2.26	0.62, 8.25	.21	1.87	0.91, 3.84	.07	1.95	0.92, 4.16	.08
2	3.05	1.71, 5.44	<.001	2.88	1.53, 5.42	<.001	5.04	1.48, 17.18	.01	2.35	1.17, 4.70	.01	1.76	0.83, 3.72	.16
3	2.61	1.44, 4.71	.001	2.48	1.30, 4.74	.005	3.61	1.03, 12.71	.04	2.49	1.23, 5.05	.008	2.57	1.22, 5.45	.02
4	2.72	1.46, 5.06	.002	2.73	1.39, 5.37	.003	5.46	1.54, 19.42	.005	3.60	1.74, 7.46	<.001	2.77	1.27, 6.05	.008

Variables	Awareness			Searching			Rating			Positive impact			Negative impact		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
5+	4.16	2.34, 7.38	<.001	3.75	2.00, 7.00	<.001	7.47	2.21, 25.27	<.001	4.07	2.05, 8.05	<.001	3.67	1.78, 7.57	<.001
Internet use			.13			.06			.17			.06			.14
> once a day	1.00			1.00			1.00			1.00			1.00		
once a day	0.72	0.47, 1.10	.09	0.63	0.40, 1.01	.05	0.59	0.30, 1.18	.06	0.63	0.38, 1.05	.06	0.61	0.34, 1.10	.07
< once a day	0.69	0.37, 1.30	.21	0.69	0.36, 1.35	.18	1.04	0.45, 2.38	.66	0.68	0.33, 1.39	.14	1.01	0.48, 2.14	.80
Health Status			.99			.99			.26			.72			.98

Table 4. Binary multivariate logistic regression analysis associated with physician-rating website relevant issues (unweighted sample).

Statistical results	Awareness	Searching	Rating	Positive impact	Negative impact
-2 Log-likelihood	1468.31	1332.10	778.44	1225.13	1040.82
Pseudo R^2 (Nagelkerke)	0.162	0.166	0.205	0.155	0.175
Constant	0.008	-20.176	-19.176	-20.011	-20.389
n	1279	1279	1279	1279	1279

Table 5. Independent factors associated with physician-rating website relevant topics (weighted sample).

Variables	Awareness			Searching			Rating			Positive impact			Negative impact		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Age			.08			.47			.99			.60			.99
Gender															
Female	1.00			1.00			1.00			1.00			1.00		
Male	0.77	0.57, 1.03	.08	0.91	0.66, 1.26	.58	1.40	0.89, 2.19	.14	0.86	0.62, 1.21	.40	0.80	0.54, 1.17	.25
Marital status															
Married	1.00			1.00			1.00			1.00			1.00		
Single	1.03	0.70, 1.50	.21	0.85	0.53, 1.34	.48	1.15	0.59, 2.22	.68	0.90	0.55, 1.48	.69	1.07	0.62, 1.84	.81
Divorced	1.02	0.63, 1.74	.96	1.05	0.58, 1.89	.89	1.56	0.71, 3.40	.27	1.00	0.52, 1.91	.99	1.13	0.54, 2.36	.75
Widowed	1.49	0.65, 3.39	.34	1.69	0.73, 3.92	.22	2.62	0.94, 7.29	.07	1.35	0.53, 3.47	.53	4.18	1.59, 10.96	.004
Education															
High School	1.00			1.00			1.00			1.00			1.00		
Technical university entrance qualification	0.65	0.41, 1.03	.06	0.72	0.44, 1.18	.19	0.96	0.49, 1.87	.90	0.94	0.57, 1.57	.82	0.94	0.54, 1.64	.84
Intermediate secondary school	0.81	0.55, 1.19	.28	0.90	0.60, 1.34	.59	0.99	0.56, 1.75	.97	1.00	0.64, 1.55	.99	0.88	0.54, 1.43	.60
Polytechnic secondary school	0.68	0.34, 1.35	.27	0.44	0.21, 0.95	.04	0.39	0.12, 1.27	.19	1.28	0.59, 2.78	.53	0.25	0.07, 0.86	.03
Secondary general school	0.76	0.45, 1.30	.32	0.45	0.24, 0.83	.01	0.87	0.41, 1.88	.73	0.84	0.43, 1.64	.61	0.62	0.29, 1.34	.22
Others	0.10	0.02, 0.63	.01	0.21	0.03, 1.31	.10	0.29	0.03, 2.54	.27	0.43	0.10, 1.84	.25	0.00	0.00,	.99
Labor															
Self-employed	1.00			1.00			1.00			1.00			1.00		
Civil servants	0.91	0.39, 2.15	.83	1.28	0.51, 3.19	.60	4.84	1.33, 17.64	.02	1.57	0.60, 4.09	.36	1.77	0.61, 5.13	.29
Employee	0.57	0.32, 1.00	.05	0.63	0.34, 1.15	.13	1.49	0.56, 3.99	.42	0.94	0.48, 1.83	.85	0.97	0.46, 2.03	.93
Apprentices	0.64	0.29, 1.42	.27	0.66	0.28, 1.54	.34	0.43	0.11, 1.64	.216	0.91	0.37, 2.21	.83	0.65	0.24, 1.78	.41
Unemployed	0.57	0.21, 1.55	.27	0.59	0.20, 1.74	.34	0.54	0.08, 3.61	.53	0.54	0.16, 1.77	.31	0.41	0.10, 1.71	.22
Pensioners	0.35	0.18, 0.70	.003	0.59	0.28, 1.23	.16	1.22	0.41, 3.58	.72	0.46	0.20, 1.10	.08	0.49	0.19, 1.26	.14
Others	0.73	0.34, 1.55	.41	0.58	0.25, 1.35	.20	2.38	0.69, 8.19	.17	0.60	0.23, 1.54	.29	0.47	0.15, 1.43	.18
Health insurance															
Statutory health insurance	1.00			1.00			1.00			1.00			1.00		
Private health insurance	0.77	0.50, 1.19	.24	0.56	0.34, 0.90	.02	0.52	0.25, 1.06	.07	0.82	0.50, 1.36	.45	0.79	0.46, 1.38	.41
Health care utilization															
No treatment	1.00		<.001	1.00		<.001	1.00		.002	1.00		<.001	1.00		<.001
1	2.32	1.17, 4.57	.02	2.02	0.96, 4.28	.07	1.82	0.51, 6.52	.36	1.88	0.82, 4.27	.13	2.05	0.83, 5.08	.12
2	3.62	1.86, 7.03	<.001	3.26	1.57, 6.77	.002	3.57	1.06, 12.00	.04	2.62	1.18, 5.81	.02	1.62	0.65, 4.01	.30
3	3.94	1.99, 7.79	<.001	3.63	1.73, 7.66	<.001	2.91	0.83, 10.25	.10	4.43	1.99, 9.86	<.001	4.25	1.75, 10.31	<.001

Variables	Awareness			Searching			Rating			Positive impact			Negative impact		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
4	3.23	1.60, 6.53	<.001	3.46	1.60, 7.49	.002	5.09	1.45, 17.86	.01	4.75	2.08, 10.86	<.001	3.90	1.56, 9.74	.004
5+	4.85	2.49, 9.43	<.001	4.10	1.98, 8.48	<.001	6.66	2.00, 22.20	.002	5.85	2.64, 12.82	<.001	5.81	2.46, 13.76	<.001
Internet use			.03			.02			.03			.010			.01
> once a day	1.00			1.00			1.00			1.00			1.00		
once a day	0.53	0.32, 0.88	.01	0.46	0.26, 0.81	.007	0.28	0.11, 0.70	.007	0.45	0.25, 0.84	.011	0.37	0.18, 0.78	.009
< once a day	0.61	0.29, 1.28	.19	0.57	0.26, 1.26	.17	1.02	0.37, 2.83	.97	0.41	0.17, 1.00	.05	0.44	0.15, 1.26	.13
Health status			.40			.14			.02			.06			.37

Table 6. Binary multivariate logistic regression analysis associated with physician-rating website relevant issues (weighted sample).

Statistical results	Awareness	Searching	Rating	Positive impact	Negative impact
-2 Log-Likelihood	1304.15	1160.57	703.95	1068.71	881.74
Pseudo R^2 (Nagelkerke)	0.276	0.268	0.294	0.259	0.278
Constant	-0.763	-20.163	-17.551	-19.926	-21.055
n	1279	1279	1279	1279	1279

Discussion

Principal Findings

Research in the field of public reporting has primarily focused on the effects of traditional instruments, which provide quality information about health care providers as related to clinical measures. However, little knowledge is available about the awareness and influence of physician-rating websites on a patient's choice. It remains unclear which key predictors are associated with the knowledge and the use of such sites. In this context, this study investigates the influence of physician-rating websites when choosing a physician and it identifies the main key predictors that are associated with the knowledge and the use of physician-rating websites by conducting a cross-sectional online survey.

In our study, approximately one-third (483/1505, 32.09%) of all respondents were aware of the existence of German physician-rating websites. This demonstrates that physician-rating websites are likely to have achieved a significant amount of publicity at least when it comes to the online population so far. Numbers from the United States indicate lower levels of awareness for such websites, although the data are older. In 2008, the Update on Consumers' Views of Patient Safety and Quality Information telephone-based survey (N=1517 respondents) showed that only 6% of Americans had heard of Hospital Compare [38], a consumer-oriented website that provides information on how well hospitals provide recommended care to their patients [39]. Another telephone survey was conducted in 2007 (N=1007 Californian adults) that showed that less than one-quarter of respondents (22% in 2007 vs 14% in 2004) had seen physician quality ratings; however, those numbers are rising [40].

In our study, one-quarter of respondents (381/1505, 25.32%) had actively searched for a physician on a German physician-rating website. Compared with other previous German surveys, this indicates an increasing amount of users on such websites. In 2011, the German Society for Consumer Research (Gesellschaft für Konsumforschung) showed a slightly lower percentage (22.6%) [41]. In 2011, another representative telephone survey of 2048 German citizens showed only 10% of respondents had searched for physicians by using German physician-rating websites (7% in 2010) [42,43]. The differences might, to a certain degree, be because of the study population (online panel vs telephone survey). In 2010, in the United States, a telephone-based survey among 3001 adults was conducted and it found that 16% of current Internet users and 19% of current online health seekers had consulted online rankings or reviews of doctors or of other providers. The same was true of another survey conducted in December 2008, which reported that 24% of respondents had used an online ranking or review when choosing a physician [44]. In general, 12% of adults have consulted online rankings or reviews of doctors or of other providers [45]. In 2011, 5% of US consumers had reported using a blog in the past year to learn about others' health care experiences (in the report, the term "blog" was not defined; thus, it remains uncertain whether these blogs are equal to physician-rating websites) [46].

Concerning our sample, 11.03% (166/1505) of respondents had posted a rating on a German physician-rating website. With reference to respondents who were exclusively aware of physician-rating websites, every third respondent had already rated a physician. The numbers observed here are higher than those from other studies. The representative telephone survey of 2048 German citizens mentioned previously showed that only 2% of respondents had posted a rating for physicians on German physician-rating websites in 2011 [42]; 1 year before

(in 2010), the number was only 1% [43]. A telephone-based US survey among 3001 adults in 2010 found that 4% of current Internet users ($n=2065$) and 6% of current online health seekers had posted an online review of a doctor. This was consistent with another study conducted in December 2008 that reported 5% [44]. Two additional studies found that 3% of adults had posted a review online about a doctor [45,46]. Therefore, only a minority have posted a rating on a physician-rating website. Rating numbers from other sectors confirm this observation (music: 5%, real estate: 4%, and cell phone: 3%) [47].

There are some surveys that investigate the impact of publicly available quality information on consumer behavior. According to our study, 65.35% (249/381) of those having performed an online search by means of a physician-rating website made their decision based on the ratings presented. This gives leverage to the statement that physicians should not underestimate the impact of such sites. Because patient awareness of such sites is likely to grow, it can be inferred that patients will be increasingly influenced by the information presented on physician-rating websites. In a US survey from 2007, it was shown that 14% of Internet users read online reviews before purchasing medical services. Of those, 76% specified that these online reviews had a significant influence on their decision [48]. However, another US study from 2008 found a much lower impact. The California telephone survey conducted in 2007 showed that only 2% of those surveyed had made a change based on information posted on a rating site (1% in 2004) [40]. Numbers from other sectors have also shown a lower impact on decision making (eg, music: 7%, cell phone: 10%) [47].

In our models, the most strongly associated variable for our physician-rating website measure was shown to be health care utilization. This is in-line with a large study conducted by Andreassen et al [49], who investigated factors that affect the health-related use of the Internet among 7 European countries. They also showed statistically significant higher odds ratios in the subsample of Internet users with higher health care utilization. Moreover, a statistically significant number of those insured by SHI were likely to be aware of such sites and use them more often when seeking and rating a physician online. Although we could not find any published evidence backing our finding, it seems probable that this is because of the fact that some large physician-rating websites are administered by SHI companies (eg, the Arztnavigator is run by the largest German SHI, Allgemeine Ortskrankenkasse). They have been promoting their website through various media channels, such as television, newspaper, radio, Internet, and membership magazines. This may have led to higher scores for those insured by an SHI company.

Higher odds ratios were calculated for female respondents in 4 variables, although only differences with respect to the awareness of the sites were proven to be statistically significant. However, significant differences could not be shown in any of our weighted models. Higher health-related online activity levels for females in general have been shown by various other studies that confirm our finding [49-57]. One explanation for that finding might be that women are more interested in health-related Internet use than are men [48,58]. Furthermore, women are more likely to register strong positive beliefs

regarding the benefits of online health searches [59]. Additionally, it seems likely that it is mostly females who take responsibility for the family's health. In cases of illness, it is mostly females who seek medical aid for themselves, their husbands, or their children [58].

In almost every model, those who were widowed were more likely to be aware of, or take advantage of, physician-rating websites, although differences were statistically significant on only 2 accounts: participants who actively searched for physicians in the unweighted models and participants who were negatively influenced in the weighted models. Because these participants had already lost a family member, it seems likely that some might have searched for health- and/or disease-related information online. Possibly, they came across physician-rating websites and were, therefore, more familiar with those websites. However, we did not find any evidence backing our assumption. In contrast, those widowed were likely to be older and possibly not familiar with online websites. Studies have shown that widowers, in general, have a lower use of eHealth [57]. Other studies have found that individuals who are married or who live with a partner are more likely to search for health information online [51,53]. We could not prove whether there were any statistically significant differences in our results regarding age in any of our models. In general, other studies (mostly telephone-based surveys) have shown that online health information seekers are relatively young (eg, [49-53,57,60,61]). However, there has not been sufficient amount of research conducted with a focus on obtaining an online sample.

No significant differences could be demonstrated with respect to education, employment, Internet use, or health status in our unweighted models. The latter is interesting because one could assume higher use of such sites with poorer health status. This assumption is backed by French and Italian evidence, which shows statistically significant higher eHealth use results for respondents with a poor perception of health or mental health as compared with those of moderate or excellent health perceptions [53,57]. However, several studies have been published showing similar results to ours. For example, Couper et al [50] demonstrated that those with better self-rated health had higher scores for health-related Internet use than those with lower health status (although not statistically significant). Andreassen et al [49] also showed an opposite impact of health status on health-related Internet use (ie, those who reported poorer health used the Internet less for health purposes). Neither Hufken and colleagues [61] nor Dumitru and colleagues [62] could prove higher health-related Internet use for those with poorer health status. However, medical indicators of health, such as a current diagnosis of long-term illness or disability, indicate a higher level of health-related use of the Internet [49]. Regarding education, our results are in-line with other studies showing that people with higher education are more likely to use the Internet for health purposes (eg, [49,53,56,57,60]). Higher results for those with higher education levels could be demonstrated in almost all models, although differences could not be proven to be statistically significant. Finally, concerning the frequency of Internet use, no statistically significant differences could be observed in the unweighted models, but they could be found in almost all of the weighted models. Higher

results for those respondents using the Internet more frequently could also be shown in other studies (eg, [53,56]).

In summary, this study demonstrated that physician-rating websites have become more common in the German online environment. Compared with previous investigations, the number of users seems to have increased. This is especially true of females with a higher education status, who are insured by SHI, and who utilize the health care sector at a higher rate. This group has demonstrated that it is aware of, and that it takes advantage of, such sites. The strongest predictor for physician-rating website use was shown to be health care utilization. Finally, it should be emphasized that physician-rating websites play an important role in choosing a physician and since their emergence in the public domain, they have influenced the decision-making process of patients. With a further increase in popularity of such sites, we predict that their influence will likely increase. Future studies are needed to investigate whether physician-rating websites have the potential to reflect the quality of care offered by health care providers.

Limitations

There are some limitations that have to be taken into account when interpreting the results of this investigation. Firstly, this

study was designed as a cross-sectional survey. Thus, we were able to identify association between exposure and outcomes. However, we could not infer cause and effect. Furthermore, the findings about the influence of physician-rating websites might differ from those studies applying an experimental design under real conditions. Therefore, we did not analyze empirical data regarding the influence in terms of numbers of encounters per quarter, the change with respect to the proportion of SHI to private health insurance patients per practice, etc. Next, we consulted an online panel for our study purposes. Obtaining an online sample instead of an offline sample meant that representation of a sample population as a whole (including online and offline samples) was not achievable. (According to the D21-Digital-Index-2013, approximately 23.5% of the German population are offline [63].) Even adjusting for differences in age, gender, education, etc, cannot compensate for the offline population. As a consequence, our findings may not be generalizable to the entire German population because the composition of the study population is predominantly middle-aged, female, and covered by private health insurance. Our study is also limited because of surveying an online panel. Those participants might be more familiar with Internet-related topics, such as searching a physician online. That could have led to higher awareness levels.

Acknowledgments

The author would like to thank Dr Phillip Goos and Ms Elke Ruppert from jameda for financing this survey.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The 28-item questionnaire.

[PDF File (Adobe PDF File), 324KB - [jmir_v15i8e187_app1.pdf](#)]

References

1. Dynan L, Goudie A, Smith RB, Fairbrother G, Simpson LA. Differences in quality of care among non-safety-net, safety-net, and children's hospitals. *Pediatrics* 2013 Feb;131(2):304-311. [doi: [10.1542/peds.2012-1089](#)] [Medline: [23296439](#)]
2. Lebrun LA, Shi L, Zhu J, Sharma R, Sripipatana A, Hayashi AS, et al. Racial/ethnic differences in clinical quality performance among health centers. *J Ambul Care Manage* 2013;36(1):24-34. [doi: [10.1097/JAC.0b013e3182473523](#)] [Medline: [23222010](#)]
3. Merchant RM, Yang L, Becker LB, Berg RA, Nadkarni V, Nichol G, American Heart Association Get With the Guideline-Resuscitation Investigators. Variability in case-mix adjusted in-hospital cardiac arrest rates. *Med Care* 2012 Feb;50(2):124-130 [FREE Full text] [doi: [10.1097/MLR.0b013e31822d5d17](#)] [Medline: [22249921](#)]
4. Tsai C, Sullivan AF, Gordon JA. Racial/ethnic differences in emergency care for joint dislocation in 53 US EDs. *Am J Emerg Med* 2012;30(9):1970-1980. [doi: [10.1016/j.ajem.2012.04.023](#)]
5. Laschet H. *ÄrzteZeitung*. 2013 Jan 21. GBA knöpft sich fünf Problem-Indikation vor URL: http://www.aerztezeitung.de/politik_gesellschaft/versorgungsforschung/article/829958/qualitaet-gba-knoepft-fuenf-problem-indikation.html [accessed 2013-08-19] [WebCite Cache ID 6J06WIDJG]
6. Porter ME, Guth C. *Chancen für das deutsche Gesundheitssystem*. Berlin: Springer/Gabler; 2012.
7. Hibbard J, Sofaer S. US Department of Health and Human Services: Agency for Healthcare Research and Quality. 2010. Best Practices in Public Reporting No 1: How To Effectively Present Health Care Performance Data To Consumers URL: <http://www.ahrq.gov/legacy/qual/pubrptguide1.htm> [accessed 2013-08-14] [WebCite Cache ID 6IrkfBQF]
8. Hibbard JH. What can we say about the impact of public reporting? Inconsistent execution yields variable results. *Ann Intern Med* 2008 Jan 15;148(2):160-161. [Medline: [18195340](#)]

9. Mukamel DB, Mushlin AI. Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State Cardiac Surgery Mortality Reports. *Med Care* 1998 Jul;36(7):945-954. [Medline: [9674613](#)]
10. Chassin MR. Achieving and sustaining improved quality: lessons from New York State and cardiac surgery. *Health Aff (Millwood)* 2002;21(4):40-51 [FREE Full text] [Medline: [12117152](#)]
11. Mukamel DB, Weimer DL, Zwanziger J, Gorthy SF, Mushlin AI. Quality report cards, selection of cardiac surgeons, and racial disparities: A study of the publication of the New York State Cardiac Surgery Reports. *Inquiry* 2005;41(4):435-446. [Medline: [15835601](#)]
12. Stevenson DG. Is a public reporting approach appropriate for nursing home care? *J Health Polit Policy Law* 2006 Aug;31(4):773-810. [doi: [10.1215/03616878-2006-003](#)] [Medline: [16971545](#)]
13. NHS. Quality and Outcomes Framework: Online GP practice results database: Search for the QOF results for your GP practice URL: <http://www.qof.ic.nhs.uk/search/index.asp> [accessed 2013-08-14] [WebCite Cache ID 6Irl5WjPE]
14. NHS. Choices: your health, your choices URL: <http://www.nhs.uk/Pages/HomePage.aspx> [accessed 2013-08-14] [WebCite Cache ID 6Irl67sNN]
15. Wubker A, Sauerland D, Wubker A. Beeinflussen bessere Qualitätsinformationen die Krankenhauswahl in Deutschland? Eine empirische Untersuchung [Does Better Quality Information Affect Hospital Choice in Germany? An Empirical Analysis. With English summary]. Stuttgart: Lucius und Lucius; 2010:467-490.
16. Geraedts M, Schwartze D, Molzahn T. Hospital quality reports in Germany: patient and physician opinion of the reported quality indicators. *BMC Health Serv Res* 2007;7:157 [FREE Full text] [doi: [10.1186/1472-6963-7-157](#)] [Medline: [17903244](#)]
17. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. *J Gen Intern Med* 2010 Sep;25(9):942-946 [FREE Full text] [doi: [10.1007/s11606-010-1383-0](#)] [Medline: [20464523](#)]
18. Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. *Medical care* 2003;41(1 Suppl):30-38. [Medline: [12544814](#)]
19. Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 2008 Jan 15;148(2):111-123. [Medline: [18195336](#)]
20. Emmert M, Gemza R, Schöffski O, Sohn S. [Public reporting in health care: the impact of publicly reported quality data on patient steering]. *Gesundheitswesen* 2012 Jun;74(6):e25-e41. [doi: [10.1055/s-0031-1285857](#)] [Medline: [21866496](#)]
21. Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. *Annu Rev Public Health* 2003;24:413-433. [doi: [10.1146/annurev.publhealth.24.100901.141005](#)] [Medline: [12428034](#)]
22. Damman OC, Hendriks M, Rademakers J, Spreeuwenberg P, Delnoij DM, Groenewegen PP. Consumers' interpretation and use of comparative information on the quality of health care: the effect of presentation approaches. *Health Expect* 2012 Jun;15(2):197-211. [doi: [10.1111/j.1369-7625.2011.00671.x](#)] [Medline: [21615637](#)]
23. Donelan K, Rogers RS, Eisenhauer A, Mort E, Agnihotri AK. Consumer comprehension of surgeon performance data for coronary bypass procedures. *Ann Thorac Surg* 2011 May;91(5):1400-5; discussion 1405. [doi: [10.1016/j.athoracsur.2011.01.019](#)] [Medline: [21524448](#)]
24. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012;14(1):e38 [FREE Full text] [doi: [10.2196/jmir.2003](#)] [Medline: [22366336](#)]
25. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. *J Med Internet Res* 2013;15(2):e24 [FREE Full text] [doi: [10.2196/jmir.2360](#)] [Medline: [23372115](#)]
26. Hardey M. Consuming professions: user-review websites and health services. *Journal of Consumer Culture* 2010;10(1):129-149. [doi: [10.1177/1469540509355023](#)]
27. Alemi F, Torii M, Clementz L, Aron DC. Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. *Qual Manag Health Care* 2012;21(1):9-19. [doi: [10.1097/QMH.0b013e3182417fc4](#)] [Medline: [22207014](#)]
28. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. *Inform Prim Care* 2009;17(4):249-253. [Medline: [20359403](#)]
29. Emmert M, Maryschok M, Eisenreich S, Schöffski O. [Websites to assess quality of care--appropriate to identify good physicians?]. *Gesundheitswesen* 2009 Apr;71(4):e18-e27. [doi: [10.1055/s-0028-1103288](#)] [Medline: [19221986](#)]
30. Kadry B, Chu LF, Kadry B, Gammars D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. *J Med Internet Res* 2011;13(4):e95. [doi: [10.2196/jmir.1960](#)] [Medline: [22088924](#)]
31. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. *J Gen Intern Med* 2012 Jun;27(6):685-692. [doi: [10.1007/s11606-011-1958-4](#)] [Medline: [22215270](#)]
32. Mostaghimi A, Crotty BH, Landon BE. The availability and nature of physician information on the internet. *J Gen Intern Med* 2010 Nov;25(11):1152-1156 [FREE Full text] [doi: [10.1007/s11606-010-1425-7](#)] [Medline: [20544300](#)]
33. Strech D, Reimann S. [German language physician rating sites]. *Gesundheitswesen* 2012 Aug;74(8-9):e61-e67. [doi: [10.1055/s-0031-1297254](#)] [Medline: [22189671](#)]

34. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. *Methods Inf Med* 2012;51(2):112-120. [doi: [10.3414/ME11-01-0045](https://doi.org/10.3414/ME11-01-0045)] [Medline: [22101427](https://pubmed.ncbi.nlm.nih.gov/22101427/)]
35. Backhaus K. *Multivariate Analysemethoden*. 13th ed. Berlin: Springer; 2011.
36. Statistisches Bundesamt. *Bevölkerung in Privathaushalten: Deutschland, Jahre, Geschlecht, Familienstand, Altersgruppen* URL: <https://www-genesis.destatis.de/genesis/online/data?operation=abrufabelleBearbeiten&levelindex=2&levelid=1366915496723&selectionname=12211-0204> [accessed 2013-08-19] [WebCite Cache ID 6J07hVKiZ]
37. The Psychiatric Research Unit. WHO-Five Well-being Index (WHO-5) URL: <http://www.who-5.org/> [accessed 2013-08-14] [WebCite Cache ID 6IrlhLED5]
38. The Henry J Kaiser Family Foundation. 2008. 2008 Update on Consumers' Views of Patient Safety and Quality Information URL: <http://www.kff.org/kaiserpolls/upload/7819.pdf> [accessed 2013-08-14] [WebCite Cache ID 6IrlsI0ow]
39. Centers for Medicare & Medicaid Services. 2013. Hospital Compare URL: <http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.html> [accessed 2013-08-14] [WebCite Cache ID 6Irlu71Qs]
40. California Healthcare Foundation. 2008. Just Looking: Consumer Use of the Internet to Manage Care URL: <http://www.chcf.org/~media/MEDIA%20LIBRARY%20Files/PDF/C/PDF%20ConsumerInternetUse.pdf> [accessed 2013-08-14] [WebCite Cache ID 6IrlumFM9]
41. Ruppert E, Jameda. 2011. GfK ermittelt: Jeder Vierte nutzt Arztempfehlungsportale URL: <http://www.jameda.de/presse/pressemeldungen/?meldung=11> [accessed 2013-08-14] [WebCite Cache ID 6IrlvLi4s]
42. National Association of Statutory Health Insurance Physicians. *Versichertenbefragung der Kassenärztlichen Bundesvereinigung 2011* URL: <http://daris.kbv.de/daris/doccontent.dll?LibraryName=EXTDARIS^DMSSLAVE&SystemType=2&LogonId=7a4437996c88e57668ad4704cbec6478&DocId=003764690&Page=1> [accessed 2013-08-19] [WebCite Cache ID 6J08AKWue]
43. National Association of Statutory Health Insurance Physicians. *Versichertenbefragung der Kassenärztlichen Bundesvereinigung 2010* URL: <http://www.kbv.de/versichertenbefragung2010.html> [accessed 2013-08-14] [WebCite Cache ID 6Irlxj0Qe]
44. Fox S, Purcell K. *Chronic Disease and the Internet*. Washington, DC: Pew Internet & American Life Project; 2010 Mar 24. URL: http://www.pewinternet.org/~media/Files/Reports/2010/PIP_Chronic_Disease.pdf [accessed 2013-08-14] [WebCite Cache ID 6IrmIcg3C]
45. Fox S. *The social life of health information, 2011*. Washington, DC: Pew Internet & American Life Project; 2011 May 12. URL: <http://www.pewinternet.org/Reports/2011/Social-Life-of-Health-Info.aspx> [accessed 2013-08-14] [WebCite Cache ID 6IrumHbEe]
46. Keckley PH. *2011 Survey of Health Care Consumers in the United States*. Washington, DC: The Deloitte Center for Health Solutions; 2011. URL: http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/US_CHS_2011ConsumerSurveyinUS_062111.pdf [accessed 2013-08-14] [WebCite Cache ID 6IrruJY0a]
47. Horrigan JB. *The Internet and consumer choice: online Americans use different search and purchase strategies for different goods*. Washington, DC: Pew Internet & American Life Project; 2008 May 18. URL: http://www.pewinternet.org/~media/Files/Reports/2008/PIP_Consumer.Decisions.pdf.pdf [accessed 2013-08-14] [WebCite Cache ID 6IrlutisBL]
48. comScore. 2007 Nov 29. *Online consumer-generated reviews have significant impact on offline purchases* URL: http://www.comscore.com/Insights/Press_Releases/2007/11/Online_Consumer_Reviews_Impact_Offline_Purchasing_Behavior [accessed 2013-08-14] [WebCite Cache ID 6IruuwhoN]
49. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. *BMC Public Health* 2007;7:53 [FREE Full text] [doi: [10.1186/1471-2458-7-53](https://doi.org/10.1186/1471-2458-7-53)] [Medline: [17425798](https://pubmed.ncbi.nlm.nih.gov/17425798/)]
50. Couper MP, Singer E, Levin CA, Fowler FJ, Fagerlin A, Zikmund-Fisher BJ. Use of the Internet and ratings of information sources for medical decisions: results from the DECISIONS survey. *Med Decis Making* 2010;30(5 Suppl):106S-114S. [doi: [10.1177/0272989X10377661](https://doi.org/10.1177/0272989X10377661)]
51. Suziedelyte A. How does searching for health information on the Internet affect individuals' demand for health care services? *Soc Sci Med* 2012 Nov;75(10):1828-1835. [doi: [10.1016/j.socscimed.2012.07.022](https://doi.org/10.1016/j.socscimed.2012.07.022)] [Medline: [22884947](https://pubmed.ncbi.nlm.nih.gov/22884947/)]
52. Miller EA, West DM. Characteristics associated with use of public and private web sites as sources of health care information: results from a national survey. *Med Care* 2007 Mar;45(3):245-251. [doi: [10.1097/01.mlr.0000244509.60556.49](https://doi.org/10.1097/01.mlr.0000244509.60556.49)] [Medline: [17304082](https://pubmed.ncbi.nlm.nih.gov/17304082/)]
53. Renahy E, Parizot I, Chauvin P. Health information seeking on the Internet: a double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. *BMC Public Health* 2008;8:69 [FREE Full text] [doi: [10.1186/1471-2458-8-69](https://doi.org/10.1186/1471-2458-8-69)] [Medline: [18291024](https://pubmed.ncbi.nlm.nih.gov/18291024/)]
54. Rice RE. Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *Int J Med Inform* 2006 Jan;75(1):8-28. [doi: [10.1016/j.ijmedinf.2005.07.032](https://doi.org/10.1016/j.ijmedinf.2005.07.032)] [Medline: [16125453](https://pubmed.ncbi.nlm.nih.gov/16125453/)]
55. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. *Int J Med Inform* 2006 Jan;75(1):29-41. [doi: [10.1016/j.ijmedinf.2005.07.029](https://doi.org/10.1016/j.ijmedinf.2005.07.029)] [Medline: [16129659](https://pubmed.ncbi.nlm.nih.gov/16129659/)]

56. 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/)]
57. Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De Vito E, et al. Surfing the internet for health information: an italian survey on use and population choices. *BMC Med Inform Decis Mak* 2011;11:21 [FREE Full text] [doi: [10.1186/1472-6947-11-21](https://doi.org/10.1186/1472-6947-11-21)] [Medline: [21470435](https://pubmed.ncbi.nlm.nih.gov/21470435/)]
58. Pirhalla R. Apotheken Umschau. 2010. Gesundheits-Managerin der Familie: Umfrage: Frauen spielen eine wichtige Rolle in der Vorsorge, aber auch die Männer sind nicht zu unterschätzen URL: <http://www.presseportal.de/pm/52678/1692696/gesundheits-managerin-der-familie-umfrage-frauen-spielen-eine-wichtige-rolle-in-der-vorsorge-aber> [accessed 2013-08-14] [WebCite Cache ID 6IrvDzUnw]
59. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. *Soc Sci Med* 2004 Nov;59(9):1795-1806. [doi: [10.1016/j.socscimed.2004.02.020](https://doi.org/10.1016/j.socscimed.2004.02.020)] [Medline: [15312915](https://pubmed.ncbi.nlm.nih.gov/15312915/)]
60. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. *Arch Intern Med* 2003 Jul 28;163(14):1727-1734. [doi: [10.1001/archinte.163.14.1727](https://doi.org/10.1001/archinte.163.14.1727)] [Medline: [12885689](https://pubmed.ncbi.nlm.nih.gov/12885689/)]
61. Hüfken V, Deutschmann M, Baehring T, Scherbaum W. Verbreitung und Sozialprofil der gesundheitsthemenbezogenen Internetnutzung: Ergebnisse einer bundesweiten Telefonumfrage. *Soz. - Präventivmed* 2004;49:381-390.
62. Dumitru RC, Bürkle T, Potapov S, Lausen B, Wiese B, Prokosch HU. Use and perception of internet for health related purposes in Germany: results of a national survey. *Int J Public Health* 2007;52(5):275-285. [Medline: [18030943](https://pubmed.ncbi.nlm.nih.gov/18030943/)]
63. Initiative D21, Infratest GmbH TNS. D21-Digital-Index: Auf dem Weg in ein digitales Deutschland?. 2013. URL: <http://www.initiatived21.de/wp-content/uploads/2013/04/digitalindex.pdf> [accessed 2013-08-14] [WebCite Cache ID 6IrvQTMZA]

Abbreviations

SHI: statutory health insurance

PHI: private health insurance

WHO: World Health Organization

WHO-5: World Health Organization Well-being Index

Edited by G Eysenbach; submitted 03.05.13; peer-reviewed by K Jethwani, S Agboola; comments to author 10.06.13; revised version received 12.06.13; accepted 14.06.13; published 28.08.13.

Please cite as:

Emmert M, Meier F, Pisch F, Sander U

Physician Choice Making and Characteristics Associated With Using Physician-Rating Websites: Cross-Sectional Study

J Med Internet Res 2013;15(8):e187

URL: <http://www.jmir.org/2013/8/e187/>

doi: [10.2196/jmir.2702](https://doi.org/10.2196/jmir.2702)

PMID: [23985220](https://pubmed.ncbi.nlm.nih.gov/23985220/)

©Martin Emmert, Florian Meier, Frank Pisch, Uwe Sander. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 28.08.2013. 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

Crowdsourcing Black Market Prices For Prescription Opioids

Nabarun Dasgupta¹, MPH, PhD; Clark Freifeld^{1,2}, MS; John S Brownstein^{1,2,3}, PhD; Christopher Mark Menone¹, MEd; Hilary L Surratt⁴, PhD; Luke Poppish⁵, MBA; Jody L Green⁵, PhD; Eric J Lavonas⁵, MD; Richard C Dart^{5,6}, MD, PhD

¹Epidemico, Boston, MA, United States

²Boston Children's Hospital, Children's Hospital Informatics Program, Harvard Medical School, Boston, MA, United States

³Boston Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, MA, United States

⁴Nova Southeastern University, Fort Lauderdale, FL, United States

⁵RADARS System, Rocky Mountain Poison and Drug Center, Denver Health, Denver, CO, United States

⁶Department of Emergency Medicine, University of Colorado School of Medicine, Aurora, CO, United States

Corresponding Author:

Nabarun Dasgupta, MPH, PhD

Epidemico

268 Newbury Street

2nd Floor

Boston, MA

United States

Phone: 1 9192603808

Fax: 1 8666571107

Email: nabarund@gmail.com

Abstract

Background: Prescription opioid diversion and abuse are major public health issues in the United States and internationally. Street prices of diverted prescription opioids can provide an indicator of drug availability, demand, and abuse potential, but these data can be difficult to collect. Crowdsourcing is a rapid and cost-effective way to gather information about sales transactions. We sought to determine whether crowdsourcing can provide accurate measurements of the street price of diverted prescription opioid medications.

Objective: To assess the possibility of crowdsourcing black market drug price data by cross-validation with law enforcement officer reports.

Methods: Using a crowdsourcing research website (StreetRx), we solicited data about the price that site visitors paid for diverted prescription opioid analgesics during the first half of 2012. These results were compared with a survey of law enforcement officers in the Researched Abuse, Diversion, and Addiction-Related Surveillance (RADARS) System, and actual transaction prices on a "dark Internet" marketplace (Silk Road). Geometric means and 95% confidence intervals were calculated for comparing prices per milligram of drug in US dollars. In a secondary analysis, we compared prices per milligram of morphine equivalent using standard equianalgesic dosing conversions.

Results: A total of 954 price reports were obtained from crowdsourcing, 737 from law enforcement, and 147 from the online marketplace. Correlations between the 3 data sources were highly linear, with Spearman rho of 0.93 ($P < .001$) between crowdsourced and law enforcement, and 0.98 ($P < .001$) between crowdsourced and online marketplace. On StreetRx, the mean prices per milligram were US\$3.29 hydromorphone, US\$2.13 buprenorphine, US\$1.57 oxycodone, US\$0.97 oxycodone, US\$0.96 methadone, US\$0.81 hydrocodone, US\$0.52 morphine, and US\$0.05 tramadol. The only significant difference between data sources was morphine, with a Drug Diversion price of US\$0.67/mg (95% CI 0.59-0.75) and a Silk Road price of US\$0.42/mg (95% CI 0.37-0.48). Street prices generally followed clinical equianalgesic potency.

Conclusions: Crowdsourced data provide a valid estimate of the street price of diverted prescription opioids. The (ostensibly free) black market was able to accurately predict the relative pharmacologic potency of opioid molecules.

(*J Med Internet Res* 2013;15(8):e178) doi:[10.2196/jmir.2810](https://doi.org/10.2196/jmir.2810)

KEYWORDS

opioids; black market; economics; drug abuse; surveillance; crowdsourcing; Internet; Silk Road; StreetRx; RADARS System; police; law enforcement

Introduction

The United States has a high level of concern with the diversion and public health consequences associated with the nonmedical use of prescription opioid analgesics [1-3]. Prescription opioid analgesics diverted from the pharmaceutical supply chain may end up being resold in open-air markets, in clubs and bars, or more subtly between friends and relatives [4].

Street price data have many applications. They are routinely collected by law enforcement agencies, which rely on accurate street prices for agents to be credible buyers or sellers in undercover operations. On a policy level, the Drug Enforcement Administration cited street price data to justify assigning buprenorphine to Schedule III, a lesser category of regulation than methadone, oxycodone, and morphine [5]. Street prices for cocaine and heroin have been used as indicators of intervention impact in public policy [6,7], including as inputs for modeling the impact of policy decisions, understanding the profits and risks in the drug trade [8,9], informing debates on prohibition [10], evaluating the impact of interdiction, informing the timing of public health efforts [11,12], and understanding the impact of globalization and economic recession on drug street prices [13-15]. Behavioral economic studies in controlled settings have been used extensively in the last decade to explain and predict human behaviors associated with addictive disorders, focusing on impulse control (discounting) and the relative likeability of substances [16,17]. In pharmaceutical development, recent guidelines for pre-approval abuse liability studies for new pharmaceuticals increase reliance on laboratory behavioral economic assessments with drug users to determine differences in willingness to pay [18], but these data from controlled settings are not connected to real-world black market street price data in post-marketing surveillance. Finally, behavioral economics research has become an important tool in understanding decision making, drug dependence, and treatment choices for a variety of reinforcing substances [19-21].

Although street price data are collected by local law enforcement, they have only occasionally been reported at a national level and are rarely made available for public health research [22]. The standard federal government source for this information, the Department of Justice's National Drug Intelligence Center (NDIC), was closed in June 2012 as part of a broader realignment of federal drug policy with no notice of future data availability.

An earlier study by our group suggested the Internet was an infrequent source of diverted drugs [23]. More recently, several anonymous online marketplaces operate via Tor hidden services (distributed traffic software enabling online anonymity) or using other identity-masking techniques. One such marketplace is Silk Road, where controlled substances can be purchased with a reasonable expectation of anonymity for both the purchaser and seller. We hypothesized that Silk Road could be a source of information on street prices for diverted prescription drugs.

Given the interest but lack of scientific efforts to collect street price information, we sought to evaluate whether online crowdsourcing could be used to measure black market street prices. Crowdsourcing is a method for harnessing distributed human intelligence, where small pieces of independently derived information are systematically collected, often using electronic tools [24]. Crowdsourcing has been used in applied biomedical research to rapidly and efficiently complete tasks that would otherwise require large amounts of time, for example to evaluate medical pictograms [25] and multilanguage patient information [26], process patient narratives [27], collect soil samples across a large area [28], identify malaria parasites in slides [29], and others [30-32]. We have previously demonstrated that electronic crowdsourcing techniques can be used to produce rapid estimates in fields as varied as infectious disease incidence and illegal wildlife trade [33-35]. We hypothesized that the same could be true for street prices. Any use of the Internet as a source of information for public health research requires careful validation against established data sources and an understanding of biases that may be present in the data. Therefore, we conducted an experiment to cross-validate 2 sources of online street price data (StreetRx and Silk Road) with a more traditional survey of law enforcement officers.

Methods

Crowdsourced Data

Launched on November 1, 2010, StreetRx is a collection of databases and websites, which gathers, organizes, and displays street price data on diverted pharmaceutical controlled substances for public health research purposes (see Figure 1) [36]. Site users anonymously submit prices they paid or heard were paid for diverted prescription drugs, specifying the drug formulation, dose, and the US city or state in which the transaction occurred. Date of entry is automatically collected. The system supports product codes and location information for the United States, Canada, and the United Kingdom. A visual photo identification feature for each formulation aids accuracy of reporting. In order to mitigate concerns about self-incrimination, the submitter can choose to identify the source of the information in 3 ways: personal experience, heard it from "someone who isn't me" (SWIM), or the Internet. Links to information on drug treatment, overdose prevention, harm reduction, safe disposal, and pain management are also provided. Site visitors can query and view submitted prices at the city level using a map interface. Users have submitted links to online pharmacies, news media, government reports, and other public sources. StreetRx averages roughly 200 visitors and 20 street price submissions per day.

Submissions for opioid analgesics that were received from the United States between January 1 and June 30, 2012, and contained data about formulation and dose strength were considered for this report. Based on previous crowdsourcing experiments, we deemed it necessary to have a systematic way

to reduce noise in the data and identify less credible submissions. Outlier prices identified by site users as “cheap” or “overpriced” on a 5-point visual analog scale were excluded; approximately one quarter of all submissions were rated in these 2 categories. Because duplicate submissions for the same drug from the same IP address less than 10 seconds apart most likely indicated submission errors, only one of the dyad was retained.

StreetRx is written in PHP programming language, with OpenLayers and jQuery user interface components. The data are stored in a MySQL relational database, on a scalable, secure hosting service with a proven track record of managing traffic spikes and high user load. Because the hosting provider also specializes in politically controversial content, the system is designed to resist attempts at being shut down due to objections to the site content. It relies on map tiles from Google Maps, but uses OpenLayers to render the map interface. The site also contains Google Analytics to track visitor volume and other statistics.

Law Enforcement Data

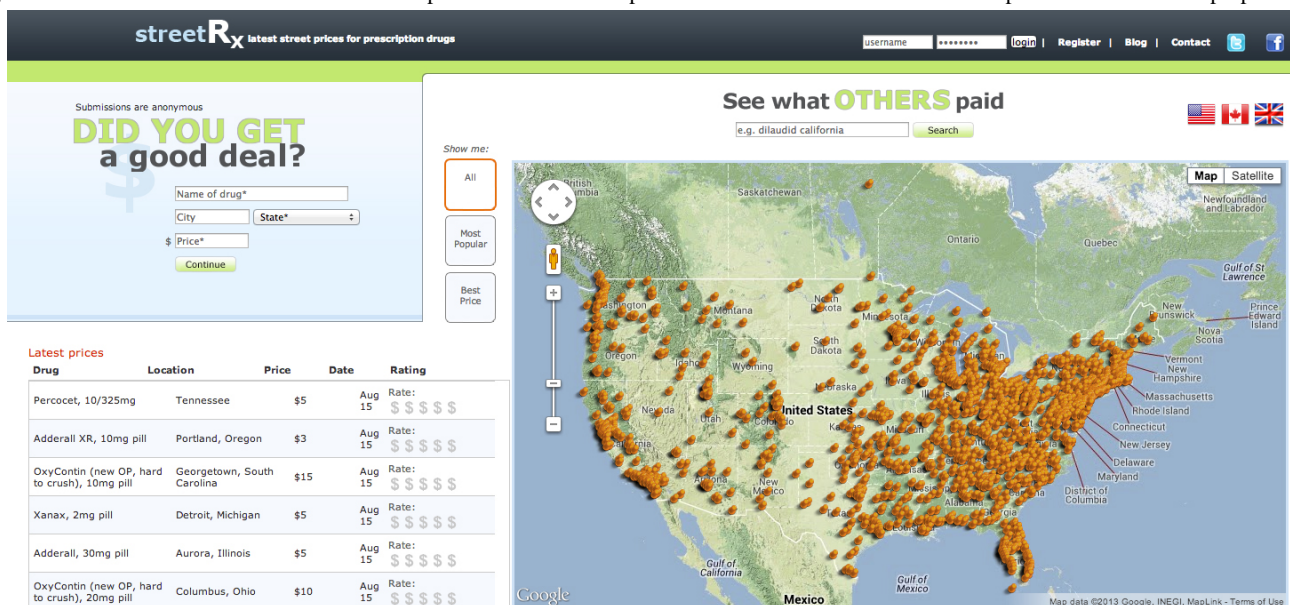
Reference data for street prices were obtained from the Researched Abuse, Diversion and Addiction Related Surveillance (RADARS) System Drug Diversion program, which collects data from approximately 280 police agencies in 49 US states on a quarterly basis. Methods of the RADARS System Drug Diversion Program, which is operated by the Center for Applied Research on Substance Use and Health Disparities, Nova Southeastern University (Miami, Florida), have been described previously [37]. For this study, a subset of 125 law enforcement reporters was selected based on prior consistency of reporting, level of diversion activity, and geographic distribution [38]. These agencies were from 46 US

states, but were not sampled in a way to make them nationally representative. A standardized electronic form was developed to collect data about the prices paid for specific drug formulations and strengths. Each reporter received the survey quarterly, in April 2012 (covering January through March 2012) and July 2012 (covering April through June 2012). Respondents were instructed to provide prices for the most common dosage strength they encountered for each opioid during the time period of interest, and not to respond for opioids with no encounters during the time period.

Dark Internet Online Marketplace Data

Silk Road is an anonymous online marketplace structured as a Tor hidden service (see Figure 2) [39]. Prospective buyers access Silk Road through a distributed network, which provides anonymity to the IP addresses of both the buyer’s Web client and the Silk Road server [40]. Silk Road uses Bitcoin (BTC), an international peer-to-peer digital currency, for payments. Bitcoin prices were converted to US dollars using the weighted average price posted on a Bitcoin exchange website on the day the sale was posted to Silk Road [41]. The exchange rate between Bitcoin and US dollars during October 2012 was approximately 11 BTC to 1 USD. A subject matter expert (author CM) manually collected (“scraped”) prices per milligram for prescription opioids in the “Opioids” section from October 1 through October 31, 2012, and collected these data on a standardized electronic data collection form. Only posts that specified that the product would be shipped from the United States were scraped. No data cleaning steps were performed and no effort was made to purchase the drugs online. No “stealth listings” (unsearchable and unlinked listings that are accessible only by buyers who have been given the URL) were scraped.

Figure 1. Screenshot of StreetRx - features street price data on diverted pharmaceutical controlled substances for public health research purposes.



StreetRx displays user-submitted information on the latest street prices for prescription drugs. StreetRx necessarily relies on user-submitted information; data should be interpreted accordingly. All submissions are anonymous. StreetRx is not affiliated with any government or law enforcement agency.
 Register | Blog | Twitter | Facebook | About | Contact us: info@streetrx.com
 External resources: Find Treatment | Overdose Prevention | Safe Pain Management | Medication Storage and Disposal | Return Unused Medications

Figure 2. Screenshot of Silk Road - an anonymous online marketplace where drugs, fireworks, and stolen goods are sold.

The screenshot shows the Silk Road website interface. At the top, there is a header with the Silk Road logo, navigation links for messages, orders, and account, and a search bar. A sidebar on the left lists various categories and their item counts. The main content area displays a list of products for sale, each with a thumbnail image, a title, a price, and an 'add to cart' button. The products shown include:

- XMAS OXYCONTIN SALE 10 x 40mg - \$235**: seller: TheDutchGuy(100), ships from: Netherlands, price: \$17.36.
- European Codeine 20mg/20 pills**: seller: FollowMe(100), ships from: Spain, price: \$1.30.
- OXYCONTIN 80MG - OP 80 Purdue plus EXTRA GIFTS**: seller: jaysonwoods, ships from: United States of America, price: \$3.25.
- 5x 8mg Subutex (buprenorphine)**: price: \$17.36.

Opioids Studied

The following drugs were initially considered for inclusion in the study: oral/sublingual dosing forms of buprenorphine, hydrocodone, hydromorphone, methadone, morphine, oxycodone, oxymorphone, tapentadol, tramadol, and transdermal patch formulations of fentanyl. Because only a small number of reports were available for tapentadol and transdermal fentanyl, these opioids were excluded from further analysis.

Data Analysis

Geometric means and 95% confidence intervals of the price per milligram were calculated for each opioid. Correlation between systems was assessed by comparing opioid-specific means using the nonparametric Spearman rank correlation coefficient (ρ), and tested with the null hypothesis that data from each pair of systems were independent with a two-tailed significance threshold of 0.05. Data management and analysis were performed using STATA version 12. In a preplanned secondary analysis, we used a standardized equianalgesic dosing conversion table [42] to convert the strength of each formulation to milligram of morphine equivalent, and compared prices on this basis. The predicted potency was calculated by dividing the mean price per milligrams for each opioid by that of morphine.

Ethics and Disclosure Statement

Law enforcement data used in this study were reviewed by the Colorado Multiple Institutional Review Board (IRB), which also provides overall ethical oversight to the RADARS System. The Drug Diversion program is classified as exempt by the Nova Southeastern University IRB, as it does not constitute human subjects research.

Results

Data from 954 StreetRx reports, 737 Drug Diversion reports, and 147 postings on Silk Road were analyzed. The most reports were received for oxycodone and hydrocodone in each system (Table 1). Correlation between the 3 data sources was high. The Spearman correlation of prices per milligram between crowdsourced and law enforcement prices was 0.93 ($P < .001$; Figure 3, top frame), and the correlation between crowdsourced and online black market prices was 0.98 ($P < .001$; Figure 3, middle frame). Data from law enforcement and the online black market were also highly correlated ($\rho = 0.90$, $P = .002$; Figure 3, bottom frame).

With the exception of morphine, there was no significant difference between the mean price per milligram of each opioid between the 3 data sources (Figure 4). The price per milligram of morphine was greater in the law enforcement survey at US\$0.67/mg (95% CI 0.59-0.75), compared with \$0.52/mg (95% CI 0.40-0.68) in the crowdsourced data and US\$0.42/mg (95% CI 0.37-0.48) for the online black market ($P = .048$).

Street prices paid for different opioids generally followed the rank order of oral equianalgesic opioid potency clinically used for rotation of opioid analgesics (Figure 4 and Table 2). The predicted potency or desirability relative to morphine was calculated by dividing the mean prices for each opioid by that of morphine. When compared to published clinical conversion guides, the predicted relative potency from crowdsourced data were similar. Oxymorphone and oxycodone had predicted potencies that were statistically indistinguishable from clinical conversion factors. Hydromorphone, hydrocodone, and methadone were valued higher on the street per milligram than in the clinic, while tramadol was valued lower on the street.

Table 1. Mean black market street prices and equianalgesic potency, US dollars per milligram, from online and law enforcement data sources, United States, 2012.

Drug	StreetRx Crowdsourcing		Drug Diversion Survey		Silk Road Marketplace	
	n	Mean, US\$ (95% CI)	n	Mean, US\$ (95% CI)	n	Mean, US\$ (95% CI)
Hydromorphone	75	3.29 (2.74-3.96)	54	4.47 (3.57-5.59)	14	3.55 (3.09-4.08)
Buprenorphine	34	2.13 (1.69-2.69)	81	2.35 (1.97-2.80)	12	2.58 (2.13-3.13)
Oxymorphone	38	1.57 (1.27-1.95)	43	1.64 (1.29-2.10)	6	1.58 (0.73-3.43)
Methadone	21	0.96 (0.71-1.29)	81	1.16 (1.01-1.37)	3	0.93 (0.65-1.34)
Oxycodone	454	0.97 (0.90-1.04)	181	0.86 (0.78-0.93)	43	0.99 (0.83-1.18)
Hydrocodone	228	0.81 (0.74-0.89)	179	0.90 (0.84-0.97)	46	0.97 (0.90-1.05)
Morphine	83	0.52 (0.40-0.68)	81	0.67 (0.59-0.75) ^a	16	0.42 (0.37-0.48) ^a
Tramadol	21	0.05 (0.03-0.07)	37	0.09 (0.07-0.12)	7	0.02 (0.01-0.03)

^aMorphine values differ between Drug Diversion Survey and Silk Road based on statistical test for possibility of random error ($P < .05$), but not between StreetRx and the other data sources.

Table 2. Mean street prices from crowdsourced data, adjusted for potency relative to morphine, United States, 2012.

Drug	Crowdsourced Data from StreetRx		Predicted Relative Potency	Clinical Equianalgesic Potency ^b
	n	Mean, US\$ (95% CI)	(95% CI) ^a	Milligrams
Hydromorphone	75	3.29 (2.74-3.96)	6.3 (5.8-6.8)	4
Oxymorphone	38	1.57 (1.27-1.95)	3.0 (2.9-3.2)	3
Methadone	21	0.96 (0.71-1.29)	1.8 (1.8-1.9)	1.5
Oxycodone	454	0.97 (0.90-1.04)	1.9 (1.5-2.2)	2
Hydrocodone	228	0.81 (0.74-0.89)	1.5 (1.3-1.8)	1
Morphine	83	0.52 (0.40-0.68)	1.0	1
Tramadol	21	0.05 (0.03-0.07)	0.1 (0.07-0.13)	0.3

^aPredicted relative potency refers to the potency or desirability as predicted by the street price relative to morphine. It was calculated by standardizing the price per milligram for each opioid against that of morphine. These numbers do not distinguish oral from other routes of administration, nor take into account time-release mechanisms. They should not be used for clinical conversion.

^bSource: United States Veterans Administration/Department of Defense Clinical Practice Guideline for the Management of Opioid Therapy for Chronic Pain, 2012.

Figure 3. Correlation between the data sources: StreetRx reports, Drug Diversion survey, and Silk Road postings.

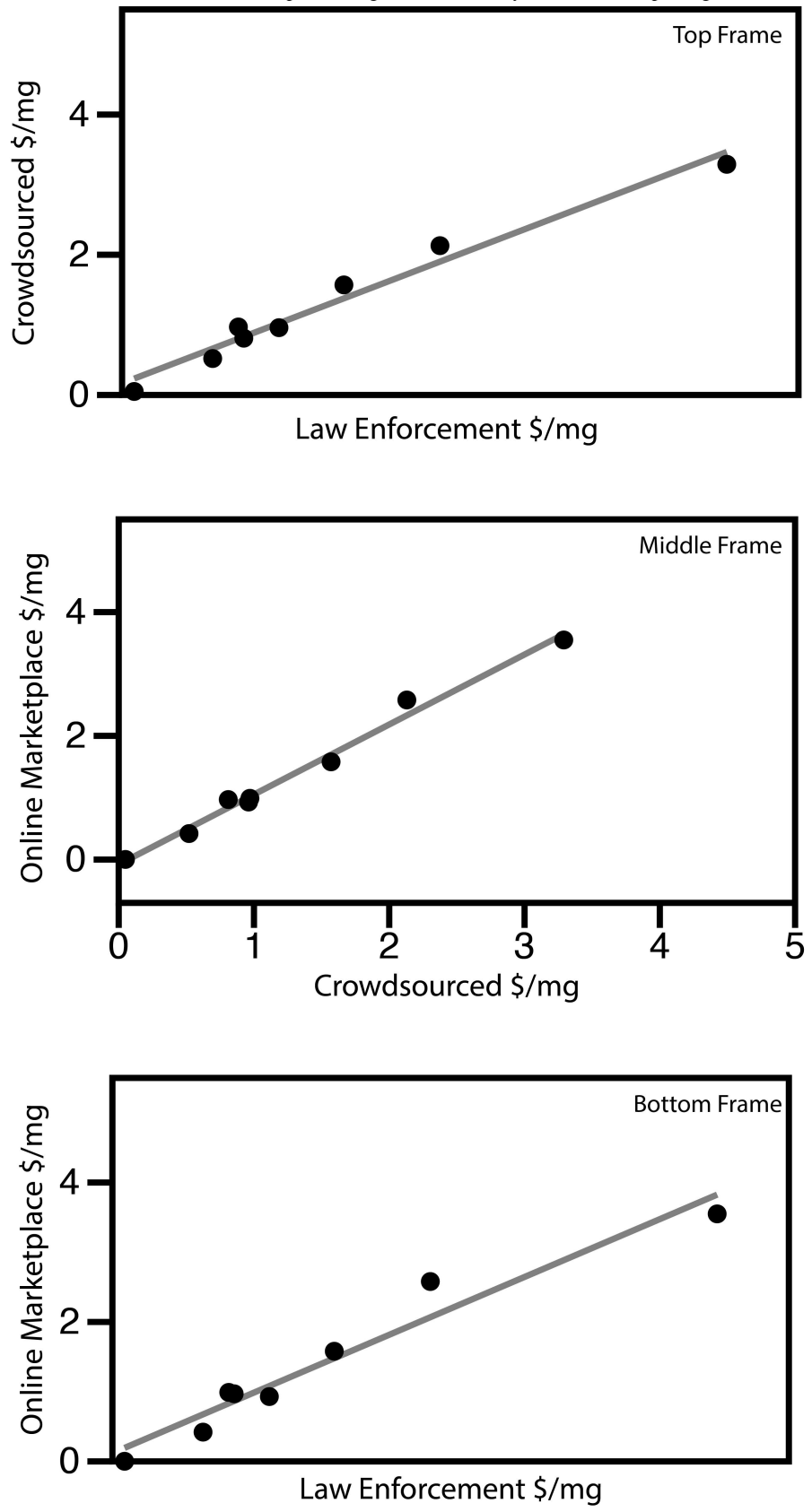
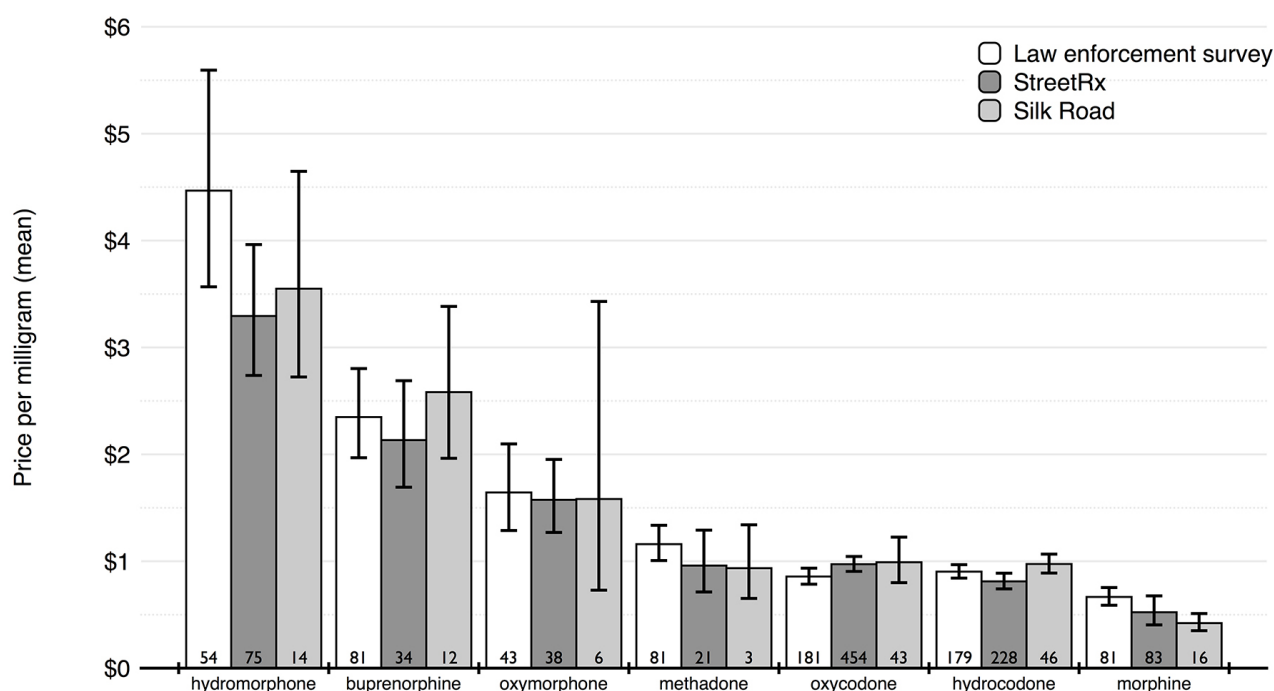


Figure 4. Mean price per milligram of each opioid analgesic, between the data sources. Numbers at the bottom of each bar indicate sample size.

Discussion

Principal Findings

Data about the street price of diverted prescription opioid medications can be useful to policymakers and public health officials, but timely and accurate data are rarely available publicly. In this paper, we present findings of a national analysis of street price data for prescription opioid analgesics. Our findings show consistent prices per milligram across 3 independent sources, and that prices for different opioid active ingredients on the black market reflect their clinically established potency. We also demonstrate the feasibility of validating crowdsourced data in the absence of a “gold standard” and document the emergence of a hidden online marketplace where drugs are sold.

Our findings are among the first to document the consistency of prices per milligram among diverted opioid analgesics. Earlier researchers, referring to heroin and cocaine, noted that “the most striking characteristics of drug prices are their high levels and extreme variability over time and space” [43]. Prescription drug prices in the United States “are affected by numerous variables, including availability, demand, law enforcement investigations, area of the country, and the relationship between the purchaser and the seller” [22]. While we observed the expected relative stability of prices during the 6 months of observation, further information is needed to understand changes over time.

For the most part, previous research has focused on the prices and purity of illicitly manufactured drugs like heroin and cocaine [13-15,43-47]. In interpreting our results, we suggest three major differences in which prescription drugs may differ from models of illicitly manufactured drugs. First, drugs are considered by economists to be “experience goods” where buyers pay before

discovering the quality of the product (such as an amusement park ride or restaurant meal). Pharmaceutical manufacturing controls create a highly uniform product where these variations are not a factor. Second, retail drug transactions for illicitly manufactured drugs often occur with standardized prices, such as “dime bags”. In these scenarios, the seller can modify the purity and weight of the product to maintain their margins. While counterfeit prescription drugs are possible, for the most part this strategy to improve profit margins is largely limited with pharmaceutical drugs. Third, urban markets for illicitly manufactured drugs are highly competitive with many small sellers creating intense competition for lower prices. Under these and other collective pressures, a seller’s reputation for quality is paramount in order to stay in business selling illicitly manufactured drugs [48]. With prescription drugs, however, these three inputs are constrained—the quality is standardized and can be assessed immediately by the buyer. Therefore, the reputation of quality of a seller may be less important than with illicit drugs. Could this create a lower threshold for initiation of selling leftover medications?

In contrast to illicitly manufactured drugs, the different prescription opioids are nearly perfect “interchangeable goods”, from the economist’s perspective (but perhaps not the pharmaceutical industry’s); it is difficult to distinguish opioids of the same potency such as heroin (diacetylmorphine) and hydromorphone. This means that we cannot look at the prices of any single prescription opioid in isolation, but must also see what is happening with the prices of other opioid molecules. We found only 2 studies that examined street prices for opioid analgesics, neither of which focused on online sources. One study found a 10x linear association between the pharmacy price and the street price of prescription opioid analgesics in Vancouver, British Columbia [49]. A study of drug users in eastern Kentucky suggested that OxyContin may serve as a

form of currency, possibly as a proxy for social capital associated with having the means to afford daily use [50].

More studies have examined Internet pharmacies. These studies concluded that the pharmacies (whether operating legally or illicitly) were found to be rarely used sources for diverted prescription drugs [4,23,51,52]. In a fundamental shift from the itinerant Internet pharmacies of the last decade, recent social network technology has allowed the emergence of online marketplaces that allow sellers to cultivate their reputation by amassing positive feedback while keeping transactions anonymous and difficult to monitor. Often the site's owners charge a commission or per-transaction fee. Monthly transactions on Silk Road have been estimated at US\$1.9 million [40], and drugs, fireworks, and stolen goods are routinely listed. Calls from US senators to shut down the sites have resulted in broad social attention [53], but the sites continue to operate. Using these sites for academic research has been limited [40], but they offer benefits. Online black market data have less opportunity for recall bias because it is possible to quantify actual transactions instead of reports of prices. However, given the moderate difficulty in accessing hidden sites relative to the open Internet, there is likely to be selection bias in terms of who uses these sites. In addition, online black markets are under pressure from authorities and may blink in and out of existence without warning, adding additional difficulty for data collection.

Websites designed for research-quality data collection via crowdsourcing and data mining are likely to cost less per report than traditional surveys and can be rapidly adapted to collect new information [54]. However, the credibility of law enforcement surveys is perceived to be higher. Other validation sources against which to compare street price data on prescription opioids are limited. With the closing of NDIC and the absence of a gold standard, we relied on triangulating street prices from 3 different sources, with our results showing remarkable consistency and robustness of prices across opioid molecules. In situations where a gold standard is not available and the behavior under question is illicit, crowdsourcing and data mining provide alternative strategies for collecting information.

Limitations

There are several limitations of this study that bear mentioning. Others have pointed out the need to take into account bulk purchasing when modeling prices of illicitly manufactured drugs

[44]. To address this, StreetRx site users were asked to note if the price they were reporting was a bulk purchase of more than 10 units. Only 10% of those submitting data positively indicated bulk purchasing; however, the question was not mandatory and we cannot exclude the possibility that others did not answer the question for separate reasons. Another limitation of this analysis was the insufficient sample size to analyze fentanyl and tapentadol data. The greatest limitation, however, lies in the sample size and non-contemporaneous data collection. Our research team was not aware of Silk Road until later in the study, but we felt that it was an important data source that was worth documenting. Geographic analyses, or at least controlling for geographic variation, were deemed to be beyond the scope of the present analysis, but an area of future research. We also did not account for the site users' history of addiction in the analyses, despite basic research suggesting that the point in the progression of addictive disease may influence willingness to pay. We hope to explore these topics in future analyses. Connecting drug prices to behaviors and health outcomes is a direction of future research.

Finally, we note that the use of equianalgesic ratios in clinical practice should be undertaken with caution, as should the interpretation of our results using these conversion numbers. In this analysis, we do not know if the opioids were diverted for self-medication, euphoria, or preventing withdrawal. The equianalgesic conversion factors were designed with opioid rotation for pain in mind, and the relative desirability for abuse or withdrawal prevention may be different. Various equianalgesic potency tables have been proposed [42,55,56], and the predicted potency conversion factors may reflect some of this variation. However, for the most part the ostensibly free black market was able to accurately predict the relatively pharmacologic potency of opioid molecules.

Conclusions

Crowdsourcing and data mining are efficient ways to collect data about street prices in an era of Internet-based social networks. These data can inform pharmacoeconomic modeling and policy analysis, and may shed light on which new controlled pharmaceutical formulations have desirability relative to others when they hit the street. While this study represents an initial foray into collecting systematic economic data for modeling black markets for prescription drugs, the methodology could be extended in the future by connecting the data to health outcomes.

Acknowledgments

The authors wish to thank Karen Lowitz for assistance in preparing this manuscript. This research was funded by the RADARS System (Denver, Colorado, USA).

Conflicts of Interest

StreetRx is entirely funded by the RADARS System (Denver, Colorado, USA), an independent nonprofit operation of the Rocky Mountain Poison and Drug Center (RMPDC), a division of Denver Health. StreetRx is operated under contract by Epidemico (Boston, Massachusetts, USA), a health data collection and analytics company. Authors of this paper include employees of the RADARS System and Epidemico. All datasets used in this publication are publicly available by contacting the RADARS System.

References

1. McDonald DC, Carlson KE. Estimating the prevalence of opioid diversion by "doctor shoppers" in the United States. *PLoS One* 2013;8(7):e69241 [FREE Full text] [doi: [10.1371/journal.pone.0069241](https://doi.org/10.1371/journal.pone.0069241)] [Medline: [23874923](https://pubmed.ncbi.nlm.nih.gov/23874923/)]
2. Berge KH, Dillon KR, Sikkink KM, Taylor TK, Lanier WL. Diversion of drugs within health care facilities, a multiple-victim crime: patterns of diversion, scope, consequences, detection, and prevention. *Mayo Clin Proc* 2012 Jul;87(7):674-682 [FREE Full text] [doi: [10.1016/j.mayocp.2012.03.013](https://doi.org/10.1016/j.mayocp.2012.03.013)] [Medline: [22766087](https://pubmed.ncbi.nlm.nih.gov/22766087/)]
3. Rigg KK, Kurtz SP, Surratt HL. Patterns of prescription medication diversion among drug dealers. *Drugs (Abingdon Engl)* 2012;19(2):144-155 [FREE Full text] [doi: [10.3109/09687637.2011.631197](https://doi.org/10.3109/09687637.2011.631197)] [Medline: [22665955](https://pubmed.ncbi.nlm.nih.gov/22665955/)]
4. Substance Abuse and Mental Health Services Administration. Results from the 2010 National Survey on Drug Use and Health: Volume I (NSDUH Series H-38A, HHS Publication No. SMA Findings). In: Summary of National Findings. Rockville, MD: Health and Human Services (Office of Applied Studies); 2010:11-4856.
5. Drug Enforcement Administration (DEA), Department of Justice. Schedules of controlled substances: rescheduling of buprenorphine from schedule V to schedule III. Final rule. *Fed Regist* 2002 Oct 7;67(194):62354-62370. [Medline: [12369590](https://pubmed.ncbi.nlm.nih.gov/12369590/)]
6. Nordt C, Stohler R. Low-threshold methadone treatment, heroin price, police activity and incidence of heroin use: the Zurich experience. *Int J Drug Policy* 2009 Nov;20(6):497-501. [doi: [10.1016/j.drugpo.2009.02.009](https://doi.org/10.1016/j.drugpo.2009.02.009)] [Medline: [19349154](https://pubmed.ncbi.nlm.nih.gov/19349154/)]
7. Kleiman MA. Enforcement swamping: A positive-feedback mechanism in rates of illicit activity. *Mathematical and Computer Modelling* 1993 Jan;17(2):65-75. [doi: [10.1016/0895-7177\(93\)90240-Y](https://doi.org/10.1016/0895-7177(93)90240-Y)]
8. Minron J. The effect of drug prohibition on drug prices: evidence from the markets for cocaine and heroin. *Review of Economics and Statistics* 2003;85(3):522-530.
9. Brown G, Silverman L. The retail price of heroin: estimation and applications. *Journal of the American Statistical Association* 1974;69(347):595-606.
10. Bretteville-Jensen AL. To legalize or not to legalize? Economic approaches to the decriminalization of drugs. *Subst Use Misuse* 2006;41(4):555-565. [doi: [10.1080/10826080500521565](https://doi.org/10.1080/10826080500521565)] [Medline: [16522563](https://pubmed.ncbi.nlm.nih.gov/16522563/)]
11. Bright DA, Ritter A. Retail price as an outcome measure for the effectiveness of drug law enforcement. *Int J Drug Policy* 2010 Sep;21(5):359-363. [doi: [10.1016/j.drugpo.2009.12.001](https://doi.org/10.1016/j.drugpo.2009.12.001)] [Medline: [20083399](https://pubmed.ncbi.nlm.nih.gov/20083399/)]
12. Caulkins JP, Tragler G, Wallner D. Optimal timing of use reduction vs. harm reduction in a drug epidemic model. *Int J Drug Policy* 2009 Nov;20(6):480-487. [doi: [10.1016/j.drugpo.2009.02.010](https://doi.org/10.1016/j.drugpo.2009.02.010)] [Medline: [19361975](https://pubmed.ncbi.nlm.nih.gov/19361975/)]
13. Costa Storti C, De Grauwe P. Globalization and the price decline of illicit drugs. *Int J Drug Policy* 2009 Jan;20(1):48-61. [doi: [10.1016/j.drugpo.2007.11.016](https://doi.org/10.1016/j.drugpo.2007.11.016)] [Medline: [18272358](https://pubmed.ncbi.nlm.nih.gov/18272358/)]
14. Caulkins JP. The global recession's effect on drug demand--diluted by inertia. *Int J Drug Policy* 2011 Sep;22(5):374-375. [doi: [10.1016/j.drugpo.2011.02.005](https://doi.org/10.1016/j.drugpo.2011.02.005)] [Medline: [21429731](https://pubmed.ncbi.nlm.nih.gov/21429731/)]
15. Bretteville-Jensen AL. Illegal drug use and the economic recession--what can we learn from the existing research? *Int J Drug Policy* 2011 Sep;22(5):353-359. [doi: [10.1016/j.drugpo.2011.03.009](https://doi.org/10.1016/j.drugpo.2011.03.009)] [Medline: [21555212](https://pubmed.ncbi.nlm.nih.gov/21555212/)]
16. Bickel WK, Miller ML, Yi R, Kowal BP, Lindquist DM, Pitcock JA. Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. *Drug Alcohol Depend* 2007 Sep;90 Suppl 1:S85-S91 [FREE Full text] [doi: [10.1016/j.drugalcdep.2006.09.016](https://doi.org/10.1016/j.drugalcdep.2006.09.016)] [Medline: [17101239](https://pubmed.ncbi.nlm.nih.gov/17101239/)]
17. Bickel WK, Marsch LA. Toward a behavioral economic understanding of drug dependence: delay discounting processes. *Addiction* 2001 Jan;96(1):73-86. [doi: [10.1080/09652140020016978](https://doi.org/10.1080/09652140020016978)] [Medline: [11177521](https://pubmed.ncbi.nlm.nih.gov/11177521/)]
18. Comer SD, Zacny JP, Dworkin RH, Turk DC, Bigelow GE, Foltin RW, et al. Core outcome measures for opioid abuse liability laboratory assessment studies in humans: IMMPACT recommendations. *Pain* 2012 Dec;153(12):2315-2324. [doi: [10.1016/j.pain.2012.07.035](https://doi.org/10.1016/j.pain.2012.07.035)] [Medline: [22998781](https://pubmed.ncbi.nlm.nih.gov/22998781/)]
19. Bickel WK, Yi R, Mueller ET, Jones BA, Christensen DR. The behavioral economics of drug dependence: towards the consilience of economics and behavioral neuroscience. *Curr Top Behav Neurosci* 2010;3:319-341. [doi: [10.1007/7854_2009_22](https://doi.org/10.1007/7854_2009_22)] [Medline: [21161759](https://pubmed.ncbi.nlm.nih.gov/21161759/)]
20. Comer SD, Bickel WK, Yi R, de Wit H, Higgins ST, Wenger GR, et al. Human behavioral pharmacology, past, present, and future: symposium presented at the 50th annual meeting of the Behavioral Pharmacology Society. *Behav Pharmacol* 2010 Jul;21(4):251-277 [FREE Full text] [doi: [10.1097/FBP.0b013e32833bb9f8](https://doi.org/10.1097/FBP.0b013e32833bb9f8)] [Medline: [20664330](https://pubmed.ncbi.nlm.nih.gov/20664330/)]
21. Bretteville-Jensen AL. Addiction and discounting. *J Health Econ* 1999 Aug;18(4):393-407. [Medline: [10539613](https://pubmed.ncbi.nlm.nih.gov/10539613/)]
22. National Drug Intelligence Center. National Prescription Drug Threat Assessment 2009. Washington, DC: US Department of Justice; 2009. URL: <http://www.justice.gov/archive/ndic/pubs33/33775/33775p.pdf> [WebCite Cache ID 6lsB86b4u]
23. Inciardi JA, Surratt HL, Cicero TJ, Rosenblum A, Ahwah C, Bailey JE, et al. Prescription drugs purchased through the Internet: who are the end users? *Drug Alcohol Depend* 2010 Jul 1;110(1-2):21-29 [FREE Full text] [doi: [10.1016/j.drugalcdep.2010.01.015](https://doi.org/10.1016/j.drugalcdep.2010.01.015)] [Medline: [20227199](https://pubmed.ncbi.nlm.nih.gov/20227199/)]
24. Good BM, Su AI. Crowdsourcing for bioinformatics. *Bioinformatics* 2013 Aug 15;29(16):1925-1933. [doi: [10.1093/bioinformatics/btt333](https://doi.org/10.1093/bioinformatics/btt333)] [Medline: [23782614](https://pubmed.ncbi.nlm.nih.gov/23782614/)]
25. Yu B, Willis M, Sun P, Wang J. Crowdsourcing participatory evaluation of medical pictograms using Amazon Mechanical Turk. *J Med Internet Res* 2013;15(6):e108 [FREE Full text] [doi: [10.2196/jmir.2513](https://doi.org/10.2196/jmir.2513)] [Medline: [23732572](https://pubmed.ncbi.nlm.nih.gov/23732572/)]
26. 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/)]

27. Maclean DL, Heer J. Identifying medical terms in patient-authored text: a crowdsourcing-based approach. *J Am Med Inform Assoc* 2013 May 5;1. [doi: [10.1136/amiainl-2012-001110](https://doi.org/10.1136/amiainl-2012-001110)] [Medline: [23645553](https://pubmed.ncbi.nlm.nih.gov/23645553/)]
28. Shelley W, Lawley R, Robinson DA. Crowd-sourced soil data for Europe. *Nature* 2013 Apr 18;496(7445):300. [doi: [10.1038/496300d](https://doi.org/10.1038/496300d)] [Medline: [23598330](https://pubmed.ncbi.nlm.nih.gov/23598330/)]
29. 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/)]
30. Ranard BL, Ha YP, Meisel ZF, Asch DA, Hill SS, Becker LB, et al. Crowdsourcing-harnessing the masses to advance health and medicine, a systematic review. *J Gen Intern Med* 2013 Jul 11;1. [doi: [10.1007/s11606-013-2536-8](https://doi.org/10.1007/s11606-013-2536-8)] [Medline: [23843021](https://pubmed.ncbi.nlm.nih.gov/23843021/)]
31. Zhai H, Lingren T, Deleger L, Li Q, Kaiser M, Stoutenborough L, et al. Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing. *J Med Internet Res* 2013;15(4):e73 [FREE Full text] [doi: [10.2196/jmir.2426](https://doi.org/10.2196/jmir.2426)] [Medline: [23548263](https://pubmed.ncbi.nlm.nih.gov/23548263/)]
32. Jarmolowicz DP, Bickel WK, Carter AE, Franck CT, Mueller ET. Using crowdsourcing to examine relations between delay and probability discounting. *Behav Processes* 2012 Nov;91(3):308-312. [doi: [10.1016/j.beproc.2012.09.001](https://doi.org/10.1016/j.beproc.2012.09.001)] [Medline: [22982370](https://pubmed.ncbi.nlm.nih.gov/22982370/)]
33. Sonricker Hansen AL, Li A, Joly D, Mekaru S, Brownstein JS. Digital surveillance: a novel approach to monitoring the illegal wildlife trade. *PLoS One* 2012;7(12):e51156 [FREE Full text] [doi: [10.1371/journal.pone.0051156](https://doi.org/10.1371/journal.pone.0051156)] [Medline: [23236444](https://pubmed.ncbi.nlm.nih.gov/23236444/)]
34. 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](https://pubmed.ncbi.nlm.nih.gov/22232449/)]
35. Chan EH, Brewer TF, Madoff LC, Pollack MP, Sonricker AL, Keller M, et al. Global capacity for emerging infectious disease detection. *Proc Natl Acad Sci U S A* 2010 Dec 14;107(50):21701-21706 [FREE Full text] [doi: [10.1073/pnas.1006219107](https://doi.org/10.1073/pnas.1006219107)] [Medline: [21115835](https://pubmed.ncbi.nlm.nih.gov/21115835/)]
36. StreetRx. URL: <http://www.streetrx.com/> [accessed 2013-08-13] [WebCite Cache ID 6lqXtcHqw]
37. Inciardi JA, Surratt HL, Cicero TJ, Kurtz SP, Martin SS, Parrino MW. The "black box" of prescription drug diversion. *J Addict Dis* 2009 Oct;28(4):332-347 [FREE Full text] [doi: [10.1080/10550880903182986](https://doi.org/10.1080/10550880903182986)] [Medline: [20155603](https://pubmed.ncbi.nlm.nih.gov/20155603/)]
38. Dart RC, Cicero TJ, Surratt HL, Rosenblum A, Bartelson BB, Adams EH. Assessment of the abuse of tapentadol immediate release: the first 24 months. *J Opioid Manag* 2012;8(6):395-402. [doi: [10.5055/jom.2012.0139](https://doi.org/10.5055/jom.2012.0139)] [Medline: [23264317](https://pubmed.ncbi.nlm.nih.gov/23264317/)]
39. Silk Road. URL: <http://silkroadvb5piz3r.onion> [accessed 2013-08-13] [WebCite Cache ID 6lqY0FGVe]
40. Christin N. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. 2012. URL: <http://arxiv.org/abs/1207.7139> [accessed 2013-08-13] [WebCite Cache ID 6lqba7V3R]
41. Mount Gox. URL: <https://www.mtgox.com/> [accessed 2013-08-13] [WebCite Cache ID 6lqY5Efja]
42. Dasgupta N, Kramer ED, Zalman MA, Carino S, Smith MY, Haddox JD, et al. Association between non-medical and prescriptive usage of opioids. *Drug Alcohol Depend* 2006 Apr 28;82(2):135-142. [doi: [10.1016/j.drugalcdep.2005.08.019](https://doi.org/10.1016/j.drugalcdep.2005.08.019)] [Medline: [16236466](https://pubmed.ncbi.nlm.nih.gov/16236466/)]
43. Caulkins J, Renter P. What price data tell us about drug markets. *Journal of Drug Issues* 1998;28(3):593-613.
44. Caulkins JP. What is the average price of an illicit drug? *Addiction* 1994 Jul;89(7):815-819. [Medline: [8081179](https://pubmed.ncbi.nlm.nih.gov/8081179/)]
45. Caulkins JP. Price and purity analysis for illicit drug: data and conceptual issues. *Drug Alcohol Depend* 2007 Sep;90 Suppl 1:S61-S68. [doi: [10.1016/j.drugalcdep.2006.08.014](https://doi.org/10.1016/j.drugalcdep.2006.08.014)] [Medline: [16971056](https://pubmed.ncbi.nlm.nih.gov/16971056/)]
46. Bach PB, Lantos J. Methadone dosing, heroin affordability, and the severity of addiction. *Am J Public Health* 1999 May;89(5):662-665. [Medline: [10224975](https://pubmed.ncbi.nlm.nih.gov/10224975/)]
47. Wood E, Tyndall MW, Spittal PM, Li K, Anis AH, Hogg RS, et al. Impact of supply-side policies for control of illicit drugs in the face of the AIDS and overdose epidemics: investigation of a massive heroin seizure. *CMAJ* 2003 Jan 21;168(2):165-169 [FREE Full text] [Medline: [12538544](https://pubmed.ncbi.nlm.nih.gov/12538544/)]
48. Caulkins J, Feichtinger G, Haunschmied J, Tragler G. Quality cycles and the strategic manipulation of value. *Operations Research* 2006;54(4):666-677.
49. Sajan A, Corneil T, Grzybowski S. The street value of prescription drugs. *CMAJ* 1998 Jul 28;159(2):139-142 [FREE Full text] [Medline: [9700324](https://pubmed.ncbi.nlm.nih.gov/9700324/)]
50. Jonas AB, Young AM, Oser CB, Leukefeld CG, Havens JR. OxyContin as currency: OxyContin use and increased social capital among rural Appalachian drug users. *Soc Sci Med* 2012 May;74(10):1602-1609. [doi: [10.1016/j.socscimed.2011.12.053](https://doi.org/10.1016/j.socscimed.2011.12.053)] [Medline: [22465379](https://pubmed.ncbi.nlm.nih.gov/22465379/)]
51. Cicero TJ, Shores CN, Paradis AG, Ellis MS. Source of drugs for prescription opioid analgesic abusers: a role for the Internet? *Pain Med* 2008 Sep;9(6):718-723. [doi: [10.1111/j.1526-4637.2007.00323.x](https://doi.org/10.1111/j.1526-4637.2007.00323.x)] [Medline: [18816331](https://pubmed.ncbi.nlm.nih.gov/18816331/)]
52. Substance Abuse and Mental Health Services Administration. Results from the 2011 National Survey on Drug Use and Health: Volume I. In: Summary of National Findings (NSDUH Series H-38A, HHS Publication No. SMA Findings). Rockville, MD: Health and Human Services (Office of Applied Studies); 2011:12-4856.

53. Chen A. The Underground Website Where You Can Buy Any Drug Imaginable. 2011. URL: <http://kotaku.com/5805928/the-underground-website-where-you-can-buy-any-drug-imaginable> [accessed 2013-08-13] [WebCite Cache ID 6lqYRQuIN]
54. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153-2157 [FREE Full text] [doi: [10.1056/NEJMp0900702](https://doi.org/10.1056/NEJMp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]
55. Haffey F, Brady RR, Maxwell S. A comparison of the reliability of smartphone apps for opioid conversion. *Drug Saf* 2013 Feb;36(2):111-117. [doi: [10.1007/s40264-013-0015-0](https://doi.org/10.1007/s40264-013-0015-0)] [Medline: [23322549](https://pubmed.ncbi.nlm.nih.gov/23322549/)]
56. Goodman LS, Gilman A, Brunton LL. Goodman & Gilman's manual of pharmacology and therapeutics. New York: McGraw-Hill Medical; 2008.

Abbreviations

IRB: Institutional Review Board

NDIC: National Drug Intelligence Center

RADARS: Researched Abuse, Diversion, and Addiction-Related Surveillance System

Edited by G Eysenbach; submitted 01.07.13; peer-reviewed by S Schnoll, Q Zhang, G Gilmore, D Ciccarone, J Zacny; comments to author 29.07.13; accepted 09.08.13; published 16.08.13.

Please cite as:

Dasgupta N, Freifeld C, Brownstein JS, Menone CM, Surratt HL, Poppish L, Green JL, Lavonas EJ, Dart RC

Crowdsourcing Black Market Prices For Prescription Opioids

J Med Internet Res 2013;15(8):e178

URL: <http://www.jmir.org/2013/8/e178/>

doi: [10.2196/jmir.2810](https://doi.org/10.2196/jmir.2810)

PMID: [23956042](https://pubmed.ncbi.nlm.nih.gov/23956042/)

©Nabarun Dasgupta, Clark Freifeld, John S Brownstein, Christopher Mark Menone, Hilary L Surratt, Luke Poppish, Jody L Green, Eric J Lavonas, Richard C Dart. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 16.08.2013. 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

Effects of Mobile Augmented Reality Learning Compared to Textbook Learning on Medical Students: Randomized Controlled Pilot Study

Urs-Vito Albrecht^{1*}, Dr med, MPH; Kristian Folta-Schoofs^{2*}, Dr rer nat; Marianne Behrends^{1*}, Dr rer biol hum; Ute von Jan^{1*}, Dr rer biol hum

¹PL Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

²Institute of Psychology, University of Hildesheim, Hildesheim, Germany

* all authors contributed equally

Corresponding Author:

Urs-Vito Albrecht, Dr med, MPH
PL Reichertz Institute for Medical Informatics
Hannover Medical School
Carl-Neuberg-Str 1
Hannover, 30625
Germany
Phone: 49 511532 ext 3508
Fax: 49 5115322517
Email: albrecht.urs-vito@mh-hannover.de

Abstract

Background: By adding new levels of experience, mobile Augmented Reality (mAR) can significantly increase the attractiveness of mobile learning applications in medical education.

Objective: To compare the impact of the heightened realism of a self-developed mAR blended learning environment (mARble) on learners to textbook material, especially for ethically sensitive subjects such as forensic medicine, while taking into account basic psychological aspects (usability and higher level of emotional involvement) as well as learning outcomes (increased learning efficiency).

Methods: A prestudy was conducted based on a convenience sample of 10 third-year medical students. The initial emotional status was captured using the "Profile of Mood States" questionnaire (POMS, German variation); previous knowledge about forensic medicine was determined using a 10-item single-choice (SC) test. During the 30-minute learning period, the students were randomized into two groups: the first group consisted of pairs of students, each equipped with one iPhone with a preinstalled copy of mARble, while the second group was provided with textbook material. Subsequently, both groups were asked to once again complete the POMS questionnaire and SC test to measure changes in emotional state and knowledge gain. Usability as well as pragmatic and hedonic qualities of the learning material was captured using AttrakDiff2 questionnaires. Data evaluation was conducted anonymously. Descriptive statistics for the score in total and the subgroups were calculated before and after the intervention. The scores of both groups were tested against each other using paired and unpaired signed-rank tests. An item analysis was performed for the SC test to objectify difficulty and selectivity.

Results: Statistically significant, the mARble group (6/10) showed greater knowledge gain than the control group (4/10) (Wilcoxon $z=2.232$, $P=.03$). The item analysis of the SC test showed a difficulty of $P=0.768$ ($s=0.09$) and a selectivity of $RPB=0.2$. For mARble, fatigue ($z=2.214$, $P=.03$) and numbness ($z=2.07$, $P=.04$) decreased with statistical significance when comparing pre- and post-tests. Vigor rose slightly, while irritability did not increase significantly. Changes in the control group were insignificant. Regarding hedonic quality (identification, stimulation, attractiveness), there were significant differences between mARble (mean 1.179, CI -0.440 to 0.440) and the book chapter (mean -0.982 , CI -0.959 to 0.959); the pragmatic quality mean only differed slightly.

Conclusions: The mARble group performed considerably better regarding learning efficiency; there are hints for activating components of the mAR concept that may serve to fascinate the participants and possibly boost interest in the topic for the remainder of the class. While the small sample size reduces our study's conclusiveness, its design seems appropriate for determining

the effects of interactive eLearning material with respect to emotions, learning efficiency, and hedonic and pragmatic qualities using a larger group.

Trial Registration: German Clinical Trial Register (DRKS), DRKS-ID: DRKS00004685; https://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00004685.

(*J Med Internet Res* 2013;15(8):e182) doi:[10.2196/jmir.2497](https://doi.org/10.2196/jmir.2497)

KEYWORDS

problem-based learning; cellular phone; education; medical; emotions

Introduction

Mobile Augmented Reality (AR) offers valuable learning opportunities and may have the potential to significantly improve the learning environment and the attractiveness of the learning process. Mobile AR blended learning environments offer a new level of experience for learners, especially in areas such as forensic medicine where ethical constraints may have to be placed on learning specific subjects in real-life scenarios. For nonmedical education, a number of studies have shown beneficial effects for AR-supported study modules. Many of these make use of AR in a mobile setting [1,2]. If used appropriately, this allows users to “immerse” themselves in the subject at hand [3] and to become involved in their own learning process.

Although there are a number of projects that integrate mobile AR for basic science education, for example, for middle-school or high-school students—and some of these also touch on subjects related to medicine [4]—projects employing such concepts for basic medical education are still rare. Regarding medicine in general, AR-based applications have so far been put to use mostly for supporting diagnostic or therapeutic purposes [5,6]. Other projects provide more or less complex simulations, such as for surgical training [7]. Although these approaches generally use Augmented Reality for complex scenarios, they all have in common that the technology is used in a stationary way that—even when used for educational purposes—keeps the users emotionally detached from the subject at hand. Often, they only serve to teach physicians about the use of specific tools, such as laparoscopic tools [7], or diagnostic methods, where a certain distance to the patient would be kept even in real-life scenarios. Although such projects certainly enhance learning by giving users experiences they would otherwise not be able to have, the aforementioned stationary AR-based diagnostic and training applications also usually do not allow full immersion of the users into the learning experience. They do not make them an integral part of the learning experience, for example, by projecting the learning content on the learner’s body and thus potentially evoking emotional responses in them that might have an additional influence on the learning process.

The current paper describes a methodological approach and study design that can be used for the purpose of measuring basic cognitive and emotional factors that must be dealt with when integrating AR-based mobile applications into medical teaching. To allow experimental testing of the aforementioned approach and study design, a mobile AR-based prototype app (mARble) was developed that can serve to provide medical students and

their educators with a versatile mobile learning environment, making it possible to simulate situations that are either ethically problematic or only rarely encountered in real life [8-11]. This prototype included content for forensic medicine. Education in this field often suffers from specific cases either being unavailable or unusable due to ethical restrictions, since—especially when dealing with survivors of a crime—additional traumatization must be strictly avoided.

In the context of this mobile learning environment, the mobile device serves to meet two basic demands for almost realistic wound pattern simulations. First, it is a portable and highly capable multimedia device, making it an ideal choice for using the technology in various learning situations. Second, through its highly advanced features, it even allows for augmented reality in these learning situations. Thus, it becomes possible to provide new, more realistic elements for the learning setting, such as the projection of wound patterns on the skin of the students, and to possibly provide a new learning experience. Using such an approach, the learners themselves become objects in their own learning process. Thus, they may more easily identify themselves with their role as a patient or an assault victim.

So far, little is known about the impact of mobile AR applications on the learner during the learning process. It is still unclear which emotions and cognitive effects are provoked in the recipient due to a higher level of realism combined with a very personal experience in a simulated setting. According to Edelman [12], emotions may have an influence on various aspects of learning. When an individual processes information, facts are attributed with “subjective significance” based on the triggered emotions and thus become a part of that individual’s interpretative system. Depending on the perceived success or failure of the learning process, for example, determined by exams, this can also have an effect on an individual’s subjective well-being [13]. In general, emotions that are evoked while learning are not only important when considering single individuals. They also have a big influence on the communication processes within groups of learners as well as with their teachers and are thus one of the key factors for overall learning success.

When taking a closer look at the significance of emotions on the learning process, a number of important questions arise: How can emotions be classified? How can their effects and benefits for the learning process be reliably quantified? In literature, there is currently no uniform scheme that sufficiently covers all of these aspects. This is additionally complicated by the fact that the impact of emotions also depends on the sociocultural context [14]. Another problem is that it is hard to

differentiate emotional aspects from related psychological concepts such as “motivation” [13]. The influence of certain emotions on the success of specific learning methods, for example, if someone is in favor of authoritative or more liberal teaching methods, may also depend on an individual’s ideological perspective [13]. When trying to describe emotions, subjective assessments must also be taken under consideration since terms such as “disgust”, “modesty”, “fear”, or “insecurity” may not always describe the exact same emotions for different individuals. It is also difficult for people to quantify their emotions exactly since emotions are regularly perceived on an instinctive, subjective, and nonverbal level.

For the purpose of the current paper, three core dimensions were identified for the evaluation of our AR-based learning environment, specifically to be able to confirm our expectations that emotional involvement during the learning process as well as learning efficiency for students learning with mobile augmented reality rise compared to those using only textbooks.

The first dimension was defined as learning effectiveness, which quantifies the influence of a learning method on the acquisition of knowledge. It was of special interest to investigate whether an improvement in knowledge is possible by means of training based on a specific learning method. The second dimension deals with the learning experience itself. This includes the usability of the provided material (practicability), the user’s identification with the learning method, the stimulation it provides, as well as its attractiveness for students and educators. The third dimension that was identified as having an influence on learning success, as indicated above, is emotion. There may be a change in the emotional status of learners after using a certain learning method. This could speak to an additional emotional involvement that may be due to the chosen learning method, for example, additional realism when using modern tools and applications that integrate augmented reality. The students become their own learning subjects, which offers a chance for experiencing an additional layer of learning: potential personal involvement.

Methods

Participants

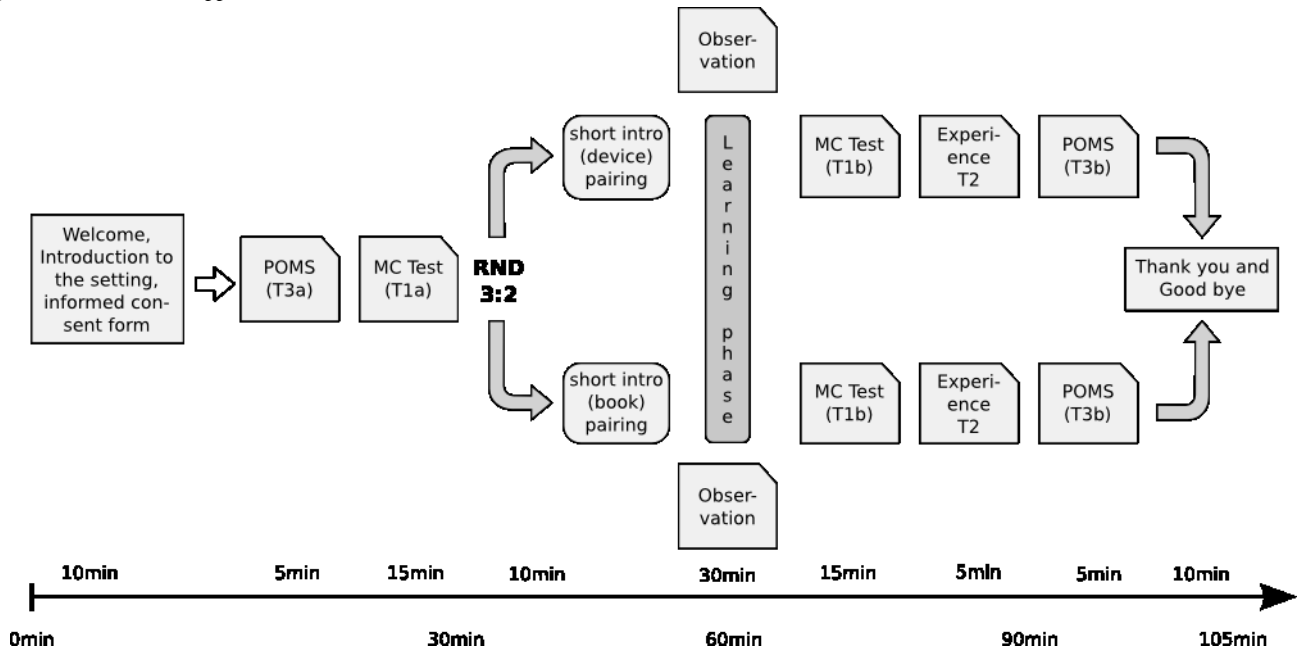
Ten third-year medical students (6 male, 4 female, mean age: 23.7 years, standard deviation: 2 years) were included in the prestudy after giving their informed consent to participate in the trial. The students had not previously participated in any regular courses dealing with the learning topic presented during the trial. Since all participants had already completed the mandatory curriculum of medical informatics, where, aside from theoretical knowledge, they were also introduced to practical aspects of using computers, it was assumed that all of them had attained at least a basic level of computer literacy.

The study was approved by the Institutional Review Board of Hannover Medical School, (ID: 1653-2012).

As shown in [Figure 1](#), to measure the emotional state of the students before the training session, all students were asked to fill out the German variation of the “Profile of Mood States” (POMS) questionnaire within a period of 5 minutes (T3a). To establish a baseline with respect to a priori knowledge of the learning topic, a 10-question standard multiple choice test about “gunshot wounds” (T1a) was given to the students, which they were asked to complete within 15 minutes. After the initial testing, the students were randomly assigned to two subgroups, named group A (6 students) and group B (4 students). Group B participated in the conservative learning situation, finding themselves in a quiet room and reading a 10-page excerpt from a standard textbook in forensic medicine [15] about “gunshot wounds”. The students were instructed to read and learn about the topic using the textbook material for a learning period of 30 minutes. While learning, they were allowed to use additional supplies such as pencils, pens, and paper to take notes and highlighters to work in the provided copies of the text material. Also, the students were free to discuss the learning topics and were specifically instructed to interact freely with other participants in their group. After 30 minutes, the students were again asked to complete the previous standard multiple choice tests comprising 10 questions (T1b). During the tests, the participants were not allowed to refer to the textbook material or their notes; they were given 15 minutes to complete the test (10 questions, 90 seconds for each answer). Afterwards, the students were asked to provide information about their learning experience (T2, 10 minutes). The POMS (T3b) questionnaire was administered to determine their emotional state after the training session. During the trial, a member of the study personnel was placed in the same room for direct observation (T4) and also to provide feedback to the students if necessary. At the end, the study participants were thanked and invited to come back on another day if they wished to try out mARble.

Group A joined the interventional arm of the trial. The group was divided into 3 pairs and each subgroup received an iPhone 4, on which the app “mARble Forensics” had already been installed, and a set of 3 paper markers. After a short greeting and introduction to the application, working with the provided markers and the learning task, the 3 pairs of students were directed into the corners of another quiet room, away from group B. The task was to learn about “gunshot wounds” using the provided iPhone and the preinstalled mARble application. The information in the mARble application contained all information relevant for later solving the multiple choice test. After 30 minutes, the students of group A had to solve the same tests as those in group B (T1b), including the multiple choice test, a questionnaire about their learning experience (T2), and the POMS questionnaire (T3b) to determine their emotional state. Just as for group B, during the trial, a member of the study personnel stayed with the students of group A for providing feedback and for direct observation (T4). After completion of all tests, participants were thanked and dismissed. The complete timeline of the individual test elements for both groups is shown in [Figure 1](#).

Figure 1. Timeline and applied tests. In the text, individual blocks are referenced via labels (T1a/b, T2, and T3a/b).



Learning Material Provided

The Application

mARble is an iOS application that was developed at the Peter L. Reichertz Institute for Medical Informatics (PLRI) at the Hannover Medical School. Using AR, virtual information can be linked to objects in the real environment, thus providing an additional layer of information to the users [10]. Code and content for mARble are kept separately. Information can easily be edited or added based on an XML-based file format. The content for the forensic module of mARble used during the course of our study was derived from and corresponded to the textbook-based learning material that was provided to the group of learners belonging to the “conventional” learning group.

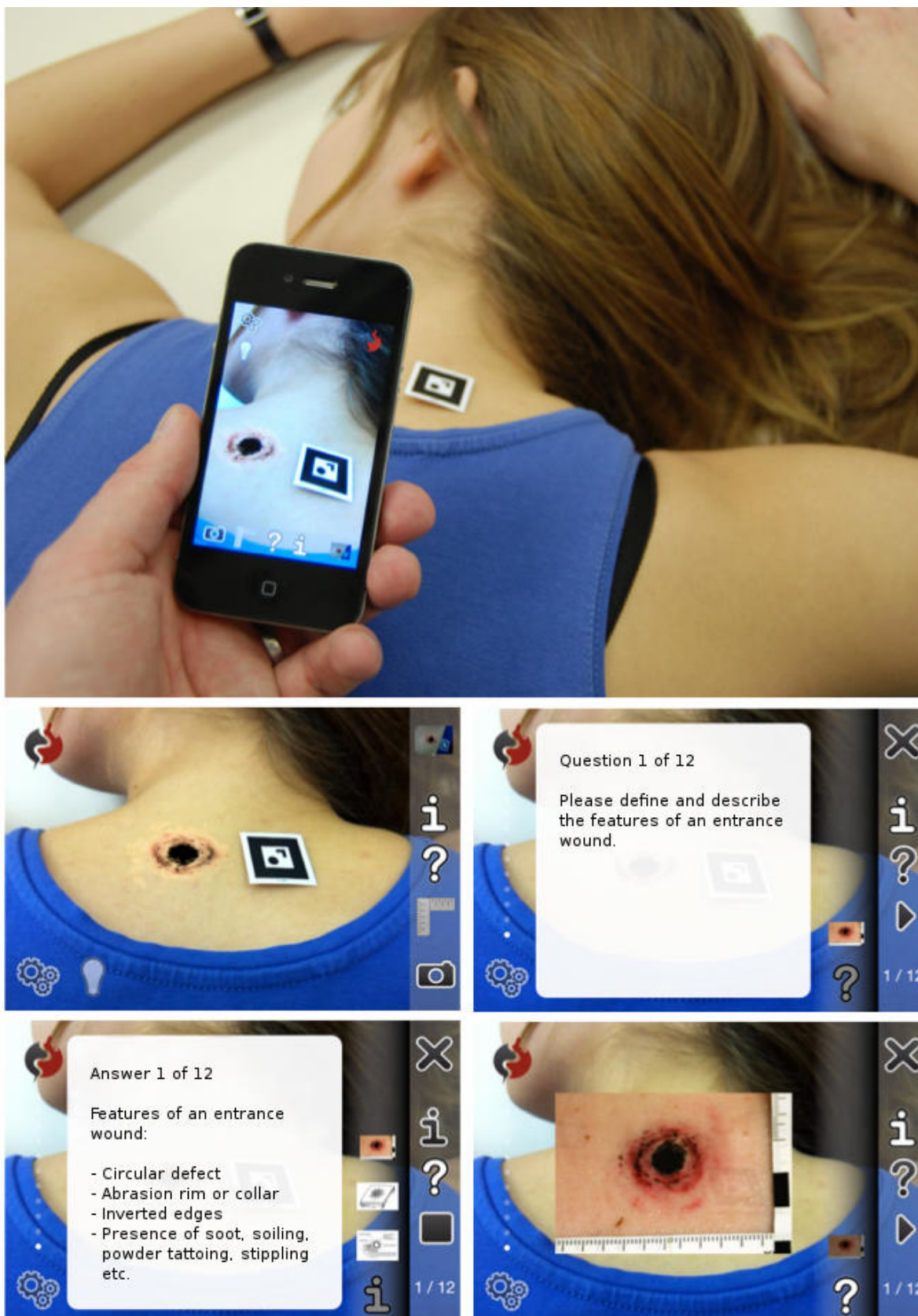
Based on this module, mARble was able to detect and interpret predefined markers representing various pathologies commonly found in forensic medicine. Each marker corresponded to a wound pattern that the students were expected to explore. When placing a marker on the student’s body, for example, on the neck, the image acquired by the iPhone’s camera was automatically overlaid with the corresponding wound pattern, such as the entrance wound of a bullet, as seen in Figure 2. Through the virtual flashcard system included in mARble, it

was also possible to view textual and multimedia background information (ie, images, drawings, video, and audio) linked to the current marker and work with the provided questions and tasks (Figure 2). Through the described approach, learners were able to construct various fictive cases by combining markers for the desired set of findings. The learning process could be documented by adding snapshots of the augmented image to a personal image gallery. Previously taken snapshots and findings could be used for review, for discussions with fellow students, or presentation purposes. When reviewing an image, it was also possible to trigger the corresponding background information as well as associated questions and tasks. Using their iPhone, students were able to examine the provided wound patterns either on themselves or on their partner; thus, they could easily immerse themselves in the learning topic.

The Conventional Learning Material

We chose the 10-page chapter “gunshot wounds” of a popular short compendium of forensic medicine in Germany [15] as learning material for group B. The textbook is very well-equipped with color and black/white pictures, schemes, and tables, as well as small repetitive summaries in colored boxes. Roughly 50% of the material consists of images and drawings.

Figure 2. The mobile Augmented Reality blended learning environment with the module “Forensic Medicine”: AR simulation of a gunshot wound and connected multimedia content.



Overall Learning Experience: Evaluation Tools

Learning Success: Multiple Choice Test (T1a, T1b)

A paper-based test consisting of 10 questions with single choice answers was used to measure the learning effectiveness. The questions and related answers were collected from a pool of material compiled by a member of the staff of the forensic medicine department. Before the trial, two members of the staff evaluated the multiple choice questionnaire with respect to comprehensibility, solvability, and time consumption. Beforehand, both the textbook extracts as well as the content provided in mARble were reviewed to determine whether the content necessary for answering all questions was sufficiently covered. Also, an item analysis before and after the learning period was conducted to take into account test difficulty (p), item discrimination (RPB), and item selectivity.

Statistical Analysis

To determine learning effectiveness (T1a and T1b), descriptive statistics were calculated, including the mean, standard deviation (SD), and mean for the score in total and the subgroups, before and after the intervention. In a noninferiority design (unpaired rank sum, Mann-Whitney U, 2-sided, Cronbach alpha=.05), the scores reached by learning with mARble were tested against the scores achieved when using the classical learning material. The calculation of T1a and T1b was based on the sum of the item values. All questionnaires were included. Those with one missing item per scale were corrected with the scale mean.

Learning Experience (T2): AttrakDiff2

Over the past few years, “user experience” has become an important factor for the acceptance of all technical innovations. Nevertheless, only vague definitions exist and there are many unanswered questions concerning the factors contributing to a good user experience [16]. Hassenzahl et al [17-19] designed a model that classifies the attributes necessary for describing products according to their pragmatic or hedonic quality and can thus be employed to describe the subjective attractiveness. This experience design concept was integrated into the test design. Following Hassenzahl’s theoretical work model, the pragmatic and hedonic qualities of an application influence a user’s subjective perception of attractiveness, resulting in respective behavioral and emotional responses. In this context, hedonic quality describes the emotional impact of a product or system, while measuring pragmatic quality offers insights into its usability or usefulness [20].

AttrakDiff2 [19] was developed as a tool by Hassenzahl’s research group to be able to quantify these qualities. The tool uses 4x7 anchor scales, in total 28 questions. The anchors are presented in the form of semantic differentials and a 7-point Likert scale is employed for rating the intensity of the items. The poles of each item are opposite adjectives (eg, “confusing-clear”, “unusual-ordinary”, “good-bad”). Each of the mean values of an item group creates a scale value for pragmatic quality (PQ), hedonic stimulation (HQ-S), hedonic identification (HQ-I), and attractiveness (ATT).

Attributes in the PQ group describe how easy the user finds it to work with the provided program or environment. Pairs of

words belonging in this group are for example “technical vs human”, “complicated vs simple”, or “impractical vs practical”.

Attributes belonging to HQ-S describe factors that encourage the personal growth of users and provide stimulation to give them the opportunity to enhance their knowledge and development. Stimulating factors can be delivered in many different ways, such as by presenting things in a novel way or by providing a new interaction style. Anchors for hedonic stimulation include attributes such as “professional vs unprofessional”, “stylish vs tacky”, or “isolating vs connective”.

The attributes falling into the HQ-I category make it possible to identify the social impact that using a product can have for users, including the “messages” that are communicated by using the evaluated product [20]. Anchors belonging in this group are, for example, “ordinary vs novel”, “conservative vs innovative”, or “undemanding vs challenging”.

Last, the attributes of the ATT group depict the overall experience a product has to offer to its users, that is, its attractiveness. Contributing attributes are, for example, “pleasant vs unpleasant”, “ugly vs attractive”, or “appealing vs unappealing”.

Statistical Methods

A Mann-Whitney U test for independent random sampling (Cronbach alpha=.05) was conducted to discriminate a possibly significant difference within the categories. Overall scores for the individual 28 attributes as well as aggregated values for each category were obtained by calculating the average values of the ratings provided by the users. To better visualize the relationship between pragmatic and hedonic qualities, the values calculated for PQ are shown on one axis and the values for HQ-I as well as HQ-S on the other axis. Combined with the confidence interval (CI) for each value, the values obtained for the ratings allow a clear differentiation between students using text-based learning or mARble.

Emotional Involvement (T3a+T3b): POMS Questionnaire, German Version

To measure the emotional state and possible psychological distress before (T3a) and after (T3b) the learning phase, we asked all students to answer the German variation of the POMS questionnaire by McNair et al [21]. This version, modified by Biehl, Dangel, and Reiser [22], consists of 35 items (adjectives) that can be divided into 4 groups describing mood disturbances, including fatigue-inertia (14 items), vigor-activity (7 items), tension-anxiety (7 items), and depression-dejection (7 items). Participants rated each item on a 7-point rating scale according to the experienced intensity (eg, “not at all”, “very little”, “a little”, “somewhat”, “fairly”, “strongly”, “very strongly”) of the corresponding mood disturbance. The triggering question was formulated as: “How do you rate your current emotional state?”

The study participants were asked to finish the survey in approximately 5 minutes. Internal consistency estimates range between Cronbach alpha=.89 and Cronbach alpha=.95 [23]. We decided to use this instrument due to its known validity and its broad usage in medical [24-26] and psychological [27] disciplines.

Direct Observation of the Participants (T4)

To be able to extensively evaluate the learning situation, observations were included in the study design. A nonparticipant observation was chosen for data collection. During the trial, the behavior of participants from both groups was observed by trained personnel. The observers, all having at least 5 years of experience teaching at a university level, were required to focus on a priori defined criteria of learning behavior as the primary basis for organizing and reporting results. Notes of additional observations of any kind were allowed and were used for further qualitative analysis. The development of observation criteria referred to statements of Schulmeister [28] on learning psychology-based factors of virtual teaching and learning, where the importance of social presence in learning settings is emphasized.

This well-known classification provided the basis for observing both groups in order to examine possible differences in participants' learning behavior. According to this classification, the following observation criteria have been selected: (1) student's communication and interactivity with peers, (2) student's focus on or distraction from the learning material, and (3) the way the student dealt with the learning object (learning material).

Results

Learning Success: Multiple Choice Test (T1a, T1b)

Comparing the results of the multiple choice tests before and after the learning period, on average, all participants showed an improvement regarding correct answers (Figure 3). Stratified for the learning method, the improvement was higher in the

mARble group with 4.7 questions (SD 2.9) compared to the control group showing an improvement of 3 questions, but also with smaller variability (SD 1.5). The difference in improvement within the mARble group was statistically significant (Wilcoxon, $z=2.232$, $P=.03$). The multiple choice test difficulty was calculated with $P=0.768$ (SD 0.09) with an item discrimination of $RPB=0.2$.

Learning Experience (T2): AttrakDiff2

Statistical analysis revealed significant differences between the mARble and the textbook groups (Table 1) for the hedonic qualities, HQ-S "stimulation" (Mann-Whitney U, $z=6.506$, $P<.001$), HQ-I "identification" (Mann-Whitney U, $z=2.825$, $P=.005$), and ATT "attractiveness" (Mann-Whitney U, $z=5.179$, $P<.001$). mARble obtained more positive ratings. The confidence interval (CI) for the hedonic quality of the mARble group was smaller than for the textbook group (Figure 4), since mARble's users were more consistent in their evaluation; therefore, mARble's ratings were applied with greater certainty. When comparing the values for pragmatic quality, the textbook group performed better than the mARble group, although this difference is not statistically significant (Mann-Whitney U, $z=-1.616$, $P=.11$). However, for the identity aspect of hedonic quality, the mARble group performed significantly better than the textbook group (Mann-Whitney U, $z=2.825$, $P=.005$). Furthermore, considering the hedonic quality "stimulation", participants in the mARble group performed much better than the textbook group (Mann-Whitney U, $z=6.506$, $P<.001$). This resulted in a difference in participants' ratings of attractiveness (Mann-Whitney U, $z=5.179$, $P<.001$), with mARble again receiving better ratings (Figures 4 and 5). Figure 6 describes the profile of the mean values and standard deviations for the word pairs stratified for the learning methods.

Figure 3. Number of incorrectly and correctly answered questions before (left) and after (right) the learning period.

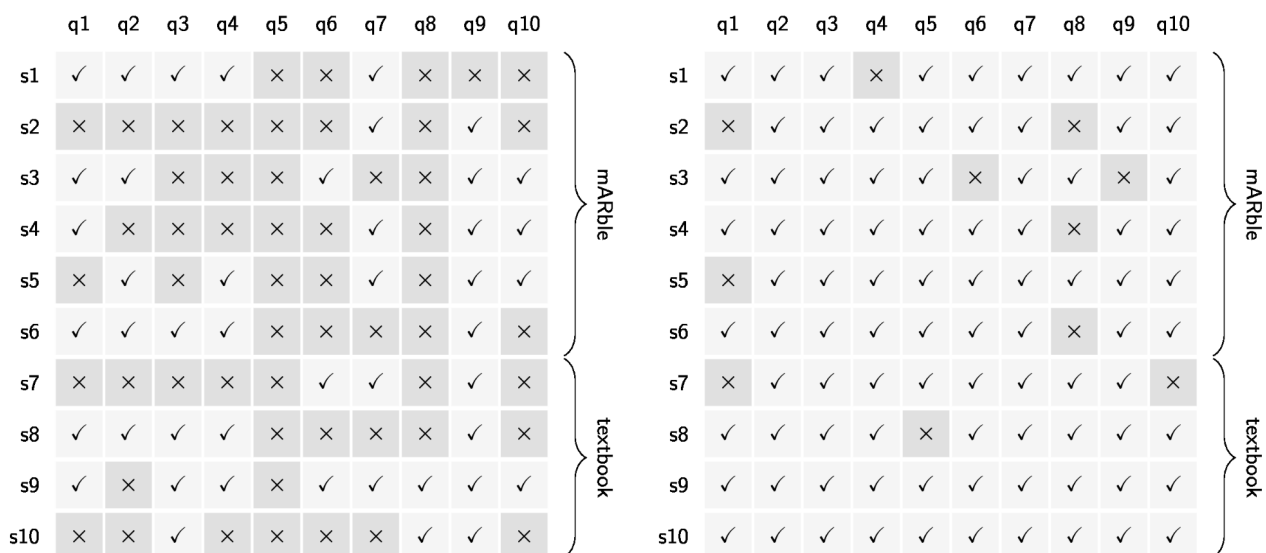


Table 1. Aggregated values calculated for the 4 qualities covered by AttrakDiff2: pragmatic quality (PQ), identification (HQ-I), stimulation (HQ-S), and attractiveness (ATT).

Group	PQ, mean (SD)	HQ-I, mean (SD)	HQ-S, mean (SD)	ATT, mean (SD)
A: mARble (n=6)	0.381 (1.168)	0.9048 (0.932)	1.452 (1.087)	1.24 (0.726)
B: textbook (n=4)	0.857 (1.758)	-0.143 (1.780)	-1.821 (1.39)	-0.57 (1.451)

Figure 4. Portfolio with average values of the dimensions PQ and HQ and the respective confidence rectangles of A (mARble) and B (textbook) on left, modified following Hassenzahl et al; corresponding values on right.

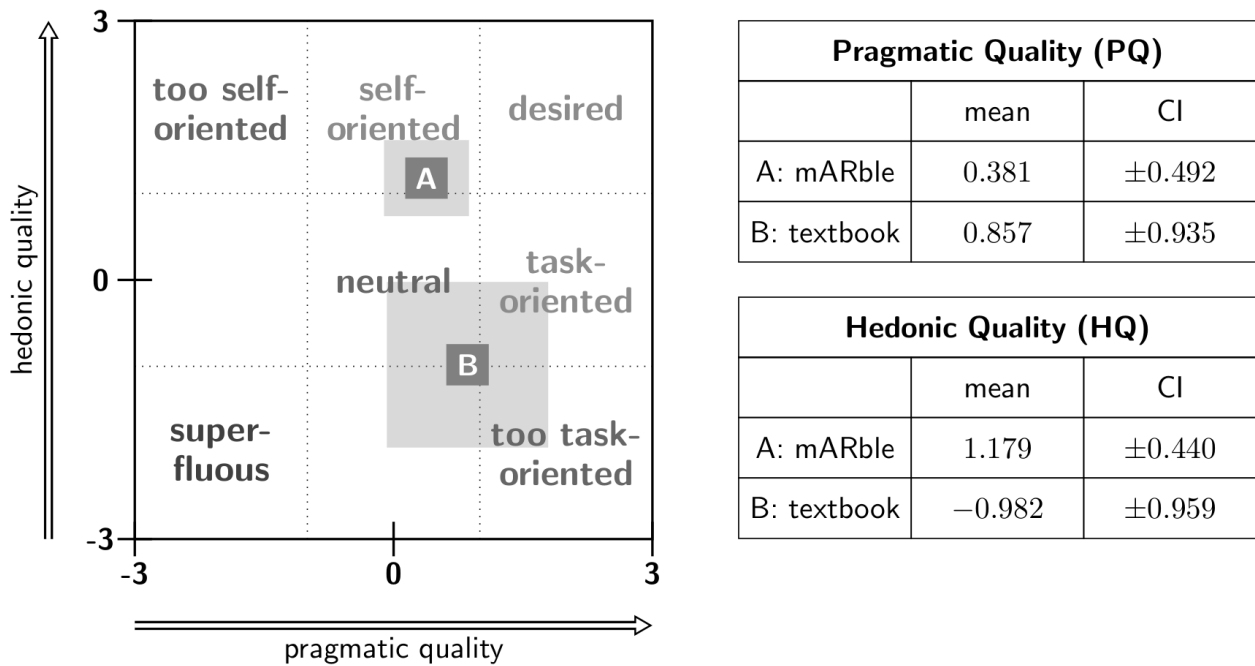


Figure 5. Average values for pragmatic quality (PQ), hedonic quality – identification (HQ-I), hedonic quality – stimulation (HQ-S), and attractiveness (ATT), based on evaluation of the AttrakDiff2 questionnaire (solid line: mARble group (6/10); dashed line: textbook group (4/10)).

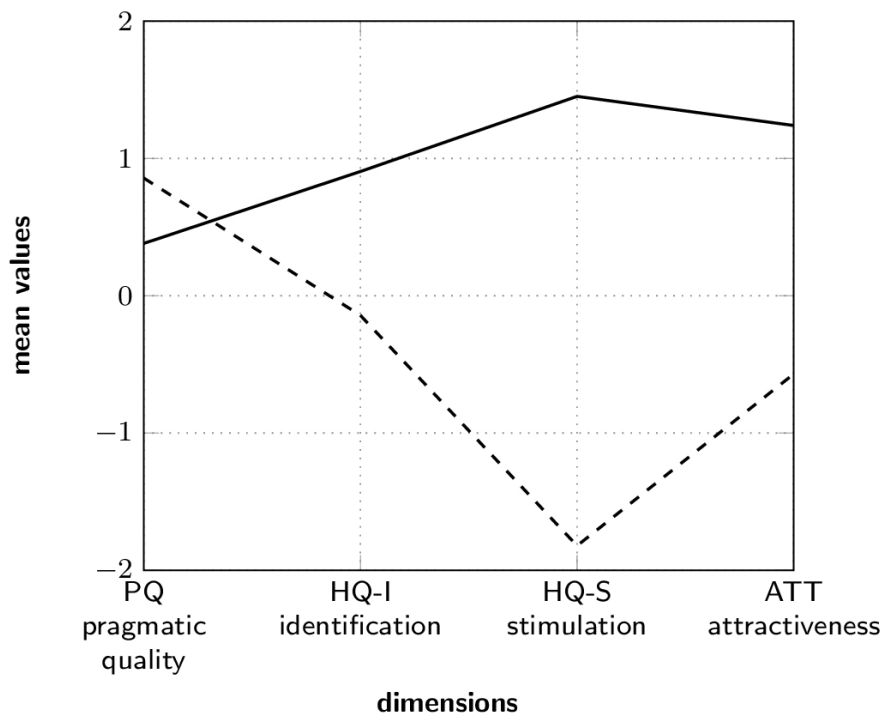
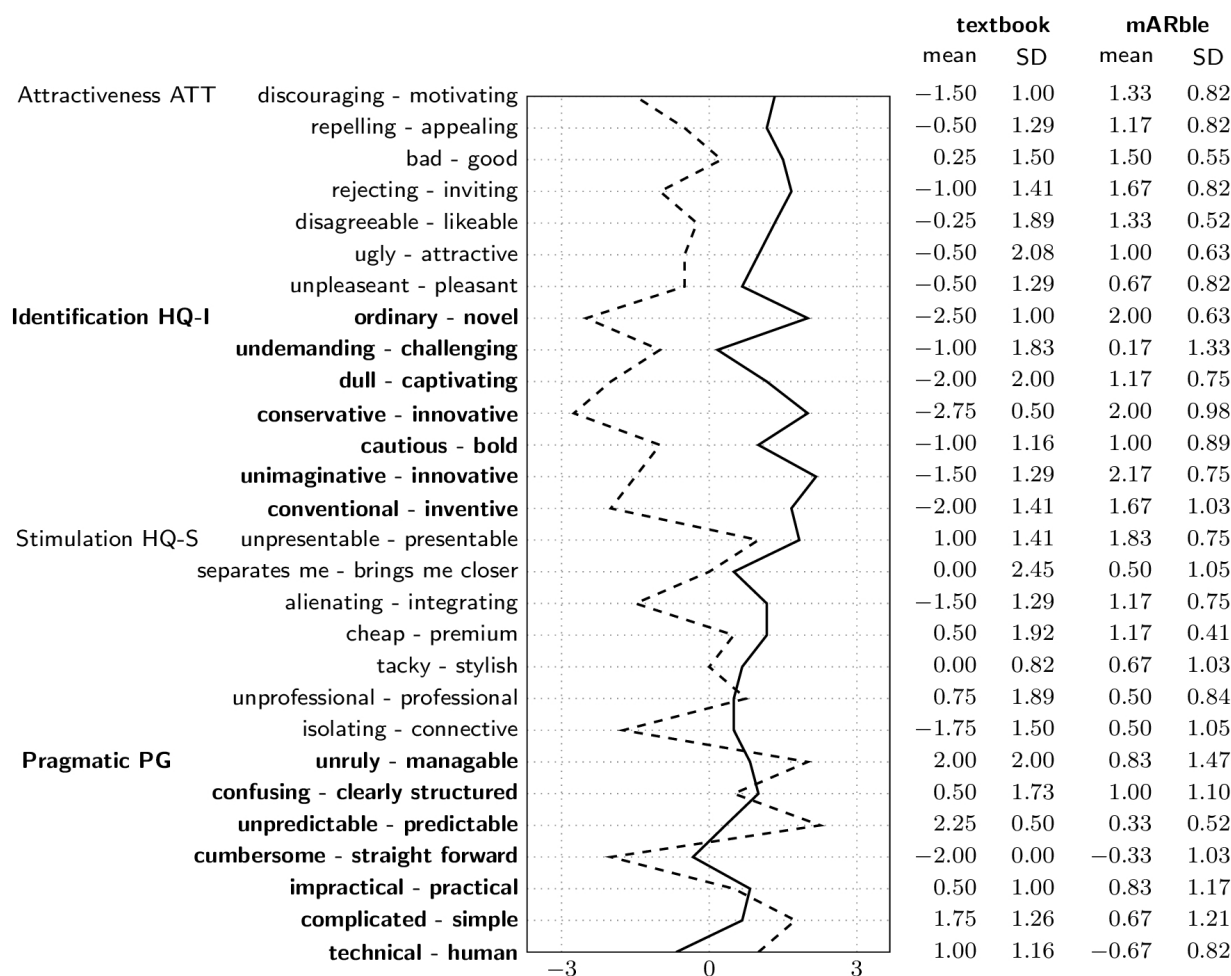


Figure 6. Description of word-pairs and calculated values: comparison between mARble (solid line) and the textbook material (dashed line).



Emotional Involvement (T3a+T3b): POMS Questionnaire, German Version

In our study, POMS was used to measure a change in the emotional state change before and after learning. The answers according to the dimensions numbness, fatigue, vigor, and irritability are shown in Table 2 and visualized in Figure 7 and were aggregated from the item values recorded for both groups of participants. A comparison of pre- and post-test values (Figure 7) showed a statistically significant decrease of fatigue ($z=2.214, P=.03$) and numbness ($z=2.07, P=.04$) for the mARble group; vigor increased slightly. Irritability did not change significantly ($z=1.166, P=.24$). The control group did not show significant changes on any of the variables.

Direct Observation of the Participants (T4)

Group A: mARble

Nonparticipant observations of the two groups showed a highly heterogeneous pattern of results. Group A (alternative learning method mARble) comprised 6 participants (P1-6) that were assigned to one of three subgroups (SG1 to SG3). In the beginning of the learning phase, all students explored the functionality of the mARble app on the iPhones. They used the marker and tried to get the picture of the linked object on their iPhone display. There was one group consisting of male

participants, one with female participants, and one mixed group. Both same-sex groups showed a high level of interaction, placing the markers on their bodies, discussing the content, and taking notes. Although the mixed group also interactively used and discussed the content, they refrained from placing the markers on their skin, instead simply placing them on the table.

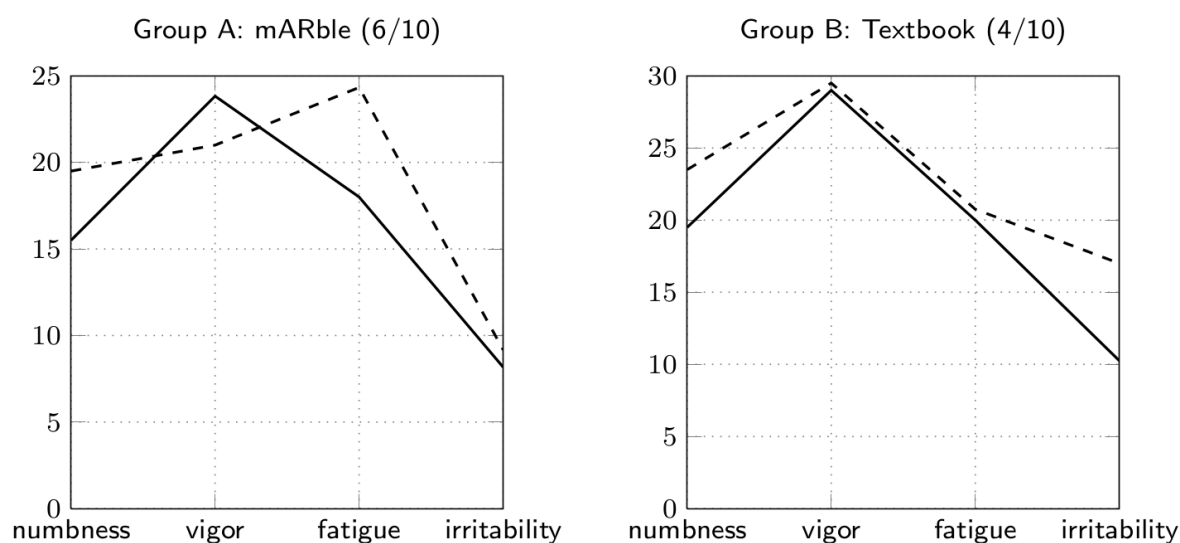
Group B: Textbook

Group B (conservative learning method) consisted of 4 participants (P1-P4) being paired and assigned to two subgroups (SG1: male, male; SG2: female, male). All participants were instructed to learn in their usual manner, but also asked to discuss the text material with their learning partner and with the whole group if they desired. Still, there were almost no dialogues with either the whole group or between the members of SG1 or SG2. From the beginning of the learning phase, all participants worked in a focused and concentrated manner, and read the text quietly. There were only very short interruptions: two participants briefly talked to each other and there was one distraction due to disruptive environmental influences. Only near the end of the learning phase was there some exchange between the participants. This did not exclusively relate to the learning content but also encompassed private matters. Further kinds of interactivity (other than communication) did not take place. The way students had worked with the text could be tracked by looking at what they had highlighted or underlined.

Table 2. Aggregated values for numbness, vigor, fatigue, and irritability.

Group	Phase	Dimensions			
		Numbness, mean (SD)	Vigor, mean (SD)	Fatigue, mean (SD)	Irritability, mean(SD)
A: mARble (n=6)					
	pre	19.5 (4.637)	21.0 (4.561)	24.33 (3.204)	9.17 (2.041)
	post	15.5 (3.209)	23.83 (9.326)	18.0 (4.147)	8.17 (1.472)
B: textbook (n=4)					
	pre	23.5 (2.887)	29.5 (9.037)	20.75 (9.570)	17.0 (5.944)
	post	19.5 (4.041)	29.0 (7.616)	20.0 (7.528)	10.25 (3.594)

Figure 7. Learning affection: pre-test (dashed line) and post-test (solid line) comparison of the aggregated values for numbness, vigor, fatigue, and irritability for both groups.



Discussion

Principal Findings

The use of mobile devices, especially when augmented reality comes into the picture, can considerably change the learning experience as well as shift it to an entirely new level [3,29], thereby providing learning experiences that are simply not possible in a conventional learning setting.

For the study, a mobile AR learning environment was developed for almost realistic wound pattern simulations in medical settings, where learners become emotionally involved in their learning process. For us, it was important to determine whether the use of AR-based solutions might trigger negative emotions or irritations in learners: when learning with mARble, by placing the markers representing specific findings on their bodies and viewing the findings on their own skin, they become emotionally involved and a part of the learning process. For the current study, we carefully investigated the emotional and cognitive impact mobile AR applications might have on the learner and the learning process in contrast to a control group of subjects that learned in a conventional medical learning environment. Since emotional reactions are often domain-specific [30], the

control group served to make sure that the measured effects were not caused by the content but by the learning medium.

Although our study showed that both learning environments induced significant cognitive improvement with respect to an increased knowledge, in direct comparison, the mARble group performed significantly better.

Regarding pragmatic quality, both methods were given average ratings. While mARble's user interface was interpreted as fairly self-oriented, the textbook was rated as rather task-oriented. Nevertheless, mARble is much more attractive to its users. The ratings also point to a significant stimulation offered by mARble. Solely in terms of pragmatic quality, the textbook material was located in the above-average region. It meets ordinary standards with regard to hedonic quality–identity. For hedonic quality–stimulation, the textbook is located in the below-average region. It does not have a stimulating effect on users. Insufficient stimulation results in a lack of motivation when using the product. Should products of similar pragmatic quality be available, users would gladly change products. The attractiveness value is located in the average region. The overall impression of the product is moderately attractive.

In comparison to the textbook material, mARble obtained better ratings with respect to vigor; its users were less fatigued after

using mARble. There was no indication of irritating properties for either of the two learning methods. Similarly, during observation, the participants showed no signs of emotional irritation. Nonetheless, the different behavior of the participants in the two groups suggests that learning with mARble and using markers on the body might provoke emotions such as shame or shyness. The participants in the gender-mixed groups did not use the markers directly on their skin but only on the table's surface. However, these aspects require further investigation.

Limitations

It might be argued whether the observed results were due to the medium, textbook vs mARble, or rather due to the chosen approach, individual learning vs social interaction. Both factors are probably closely interrelated. For example, for the control group learning with the textbook material, discussion and interaction between the participants was not prohibited in any way; rather, students were specifically asked to interact with other members of their learning group if they desired. Although they sat close to each other, they chose individual learning instead of social interaction. However, due to the limited number of participants in this prestudy, we refrained from specifically using separate groups for testing both textbook and mARble in individual learning sessions as well as in an interactive way, and instead let the participants choose what suited them best.

Nevertheless, the effects of both learning approaches cannot be completely separated from the effects of the chosen learning material. Even if students learning with textbooks were to choose social interaction during their learning phase, they would still need time periods to read the material. On the other hand, with AR-based approaches, they can collaboratively use the presented material by listening to the content, looking at overlaid images and additional material right away, which encourages social interaction. Still, the possible bias of the results we

achieved by direct comparison between the textbook-based learning, which is assumed to be an individual learning method when the students are not explicitly ordered to collaborate vs the AR-based learning, where we expected a more collaborative learning process by deploying one device to each pair of students, remains a limitation of the presented study. An upcoming study should consider this and take the effects of social interaction or missing interaction during the learning process into account, for example by using individual as well as interactive setups for both textbook-based and AR-based learning and comparing the results not only between the learning methods themselves but also between collaborative as well as individual learning settings for both methods.

Additionally, the multiple choice questions used for rating the students' performance should be selected more carefully. When looking at the results of the item analysis of the multiple choice questions, it became clear that the item difficulty and item discrimination of the questions used during the presented study still had room for improvement. This could, for example, be alleviated by selecting items where the values for these criteria are already known, such as by choosing questions from previously conducted official exams rather than self-developed items.

Conclusions

Although limited by the small sample size, and the limited amount of time and content available, the chosen evaluation setup allowed for certain conclusions regarding the desired factors; a study using a larger group of participants, based on our current study's design, may provide more conclusive results regarding various aspects of interactive mobile learning tools such as mARble in comparison to conventional learning material.

Acknowledgments

We would like to thank Bernhard Häussermann and Felix Strübing for technical support during the development of mARble, Anke Mittelstädt, Christoph Noll for assistance with the photography, and Dr. Jörn Krückeberg for stimulating discussions. Special thanks also go to Prof Dr Herbert Matthies (PLRI) and Prof Dr Michael Klintschar, Institute of Legal Medicine, Hannover Medical School for their effort and expertise while conducting this study. We also acknowledge support by Deutsche Forschungsgemeinschaft for the publication costs. The manuscript was checked against the CONSORT-EHEALTH checklist [31].

Conflicts of Interest

None declared.

References

1. Parhizkar B, Gebriil ZM, Obeidy WK, Ngan MNA, Chowdhury SA, Lashkari AH. Android mobile augmented reality application based on different learning theories for primary school children. In: Proceedings of the International Conference on Multimedia Computing and Systems. 2012 Presented at: Multimedia Computing and Systems (ICMCS), International Conference; May 10-12, 2012; Tangier, Morocco p. 404-408. [doi: [10.1109/ICMCS.2012.6320114](https://doi.org/10.1109/ICMCS.2012.6320114)]
2. Larsen YC, Buchholz H, Brosda C, Bogner FX. Evaluation of a portable and interactive augmented reality learning system by teachers and students. In: Augmented Reality in Education. 2011 Presented at: EDEN - 2011 Open Classroom Conference; October 27 - 29, 2011; Ellinogermaniki Agogi, Athens, Greece p. 41-50 URL: http://www.sctg.eu/materials/sctgo_proceedings_low.pdf

3. Dede C. Immersive interfaces for engagement and learning. *Science* 2009 Jan 2;323(5910):66-69. [doi: [10.1126/science.1167311](https://doi.org/10.1126/science.1167311)] [Medline: [19119219](https://pubmed.ncbi.nlm.nih.gov/19119219/)]
4. Vilkoniene M. Influence of augmented reality technology upon pupils' knowledge about human digestive system: The results of the experiment. *US-China Education Review* 2009;6(1):36-43.
5. Weidert S, Wang L, von der Heide A, Navab N, Euler E. [Intraoperative augmented reality visualization. Current state of development and initial experiences with the CamC]. *Unfallchirurg* 2012 Mar;115(3):209-213. [doi: [10.1007/s00113-011-2121-8](https://doi.org/10.1007/s00113-011-2121-8)] [Medline: [22406917](https://pubmed.ncbi.nlm.nih.gov/22406917/)]
6. Kersten-Oertel M, Chen SS, Drouin S, Sinclair DS, Collins DL. Augmented reality visualization for guidance in neurovascular surgery. *Stud Health Technol Inform* 2012;173:225-229. [Medline: [22356991](https://pubmed.ncbi.nlm.nih.gov/22356991/)]
7. Nugent E, Shirilla N, Hafeez A, O'Riordain DS, Traynor O, Harrison AM, et al. Development and evaluation of a simulator-based laparoscopic training program for surgical novices. *Surg Endosc* 2013 Jan;27(1):214-221. [doi: [10.1007/s00464-012-2423-0](https://doi.org/10.1007/s00464-012-2423-0)] [Medline: [22773232](https://pubmed.ncbi.nlm.nih.gov/22773232/)]
8. von Jan U, Noll C, Behrends M, Albrecht UV. mARble - Augmented Reality in Medical Education. *Biomed Tech (Berl)* 2012;57(Suppl. 1):67-70. [doi: [10.1515/bmt-2012-4252](https://doi.org/10.1515/bmt-2012-4252)] [Medline: [22944785](https://pubmed.ncbi.nlm.nih.gov/22944785/)]
9. Albrecht U, von Jan U, Häussermann B, Krückeberg J, Matthies H. Hands-on mobile Augmented Reality simulations to learn about ethically sensitive topics in medicine through self-experience. Schooling of awareness and understanding. In: *Proceedings of the IADIS International Conference Mobile Learning. IADIS Press; 2012 Presented at: IADIS International Conference Mobile Learning; March 11-13, 2012; Berlin, Germany p. 447-452.*
10. Albrecht UV, von Jan U, Krückeberg J, Behrends M, Matthies HK. Medical Students Experience the Mobile Augmented Reality Blended Learning Environment mARble – An attractive concept for the Net-generation? In: *Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA 2011). IADIS Press; 2011 Presented at: IADIS International Conference on Cognition and Exploratory Learning in the Digital Age; March 10-12, 2011; Avila, Spain p. 263-266.*
11. Noll C, von Jan U, Schaft T, Matthies HK, Albrecht UV. Sehen, Fühlen, Erfahren – Rechtsmedizin mit mARble erleben. In: *Reiterer H, Deussen O, editors. Mensch & Computer 2012 – Workshopband: Interaktiv informiert – allgegenwärtig und allumfassend!?. Munich: Oldenbourg Verlag; 2012:241-246.*
12. Edelmann W. *Lernpsychologie*. 6th ed. Weinheim: Psychologie Verlags Union; 2000.
13. Götz T, Zirngibl A, Pekrun R, Hall N. Emotions, learning and achievement from an educational-psychological perspective. In: *Mayring P, von Rhoeneck C, editors. Learning emotions: the influence of affective factors on classroom learning. Frankfurt: Peter Lang; 2003:9-28.*
14. O'Regan K. Emotion and eLearning. *Journal of Asynchronous Learning Networks* 2003 Sep;7(3):78-92.
15. Madea B. *Praxis Rechtsmedizin: Befunderhebung, Rekonstruktion, Begutachtung*. Heidelberg: Springer; 2006.
16. Isleifsdóttir J, Lárusdóttir M. Measuring the User Experience of a Task Oriented Software. In: *Proceedings of the International Workshop on Meaningful Measures. Toulouse, France: Institute of Research in Informatics of Toulouse (IRIT); 2008 Presented at: International Workshop on Meaningful Measures; June 18, 2008; Reykjavik p. 27-31 URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.177.7123&rep=rep1&type=pdf#page=29>*
17. Hassenzahl M, Burmester M, Koller F. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: *Ziegler J, Szwillus G, editors. Mensch & Computer 2003. Interaktion in Bewegung. Stuttgart: B.G. Teubner; 2003:187-196.*
18. Hassenzahl M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Comp. Interaction* 2004 Dec 1;19(4):319-349. [doi: [10.1207/s15327051hci1904_2](https://doi.org/10.1207/s15327051hci1904_2)]
19. Hassenzahl M, Burmester M, Koller F. Der User Experience (UX) auf der Spur: Zum Einsatz von www.attrakdiff.de. In: *Brau H, Diefenbach S, Hassenzahl M, Koller F, Peissner M, Röse K, editors. Usability Professionals. Stuttgart: German Chapter der Usability Professionals Association; 2008:78-82.*
20. Hartson R, Pardha P. In: *Roumeliotis R, Bevans D, editors. The UX Book: Process and Guidelines for Ensuring a Quality User Experience. Waltham, MA, USA: Morgan Kaufmann (Elsevier); 2012.*
21. McNair DM, Lorr M, Droppleman LF. *Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Services; 1971.
22. Biehl B, Dangel S, Reiser A. POMS: Profile of Mood States – deutsche Fassung. In: *Collegium Internationale Psychiatriae Scalarum (CIPS), editor. Internationale Skalen für Psychiatrie. Weinheim: Beltz Test Gesellschaft; 1981.*
23. Albani C, Blaser G, Geyer M, Schmutzer G, Brähler E, Bailer H, et al. [The German short version of "Profile of Mood States" (POMS): psychometric evaluation in a representative sample]. *Psychother Psychosom Med Psychol* 2005 Jul;55(7):324-330. [doi: [10.1055/s-2004-834727](https://doi.org/10.1055/s-2004-834727)] [Medline: [15986282](https://pubmed.ncbi.nlm.nih.gov/15986282/)]
24. Pesudovs K, Weisinger HS, Coster DJ. Cataract surgery and changes in quality of life measures. *Clin Exp Optom* 2003 Jan;86(1):34-41. [Medline: [12568649](https://pubmed.ncbi.nlm.nih.gov/12568649/)]
25. Cramer JA, Hammer AE, Kustra RP. Improved mood states with lamotrigine in patients with epilepsy. *Epilepsy Behav* 2004 Oct;5(5):702-707. [doi: [10.1016/j.yebeh.2004.07.005](https://doi.org/10.1016/j.yebeh.2004.07.005)] [Medline: [15380122](https://pubmed.ncbi.nlm.nih.gov/15380122/)]
26. Crumlish CM. Coping and emotional response in cardiac surgery patients. *West J Nurs Res* 1994 Feb;16(1):57-68. [Medline: [8128669](https://pubmed.ncbi.nlm.nih.gov/8128669/)]

27. Dolbeault S, Cayrou S, Brédart A, Viala AL, Desclaux B, Saltel P, et al. The effectiveness of a psycho-educational group after early-stage breast cancer treatment: results of a randomized French study. *Psychooncology* 2009 Jun;18(6):647-656. [doi: [10.1002/pon.1440](https://doi.org/10.1002/pon.1440)] [Medline: [19039808](https://pubmed.ncbi.nlm.nih.gov/19039808/)]
28. Schulmeister R. Didaktisches Design aus hochschuldidaktischer Sicht – Ein Plädoyer für offene Lernsituationen. In: Rinn U, Meister DM, editors. *Didaktik und Neue Medien. Konzepte und Anwendungen in der Hochschule*. Münster: Waxmann; 2004.
29. Dede C, Barab S. Emerging Technologies for Learning Science: A Time of Rapid Advances. *J Sci Educ Technol* 2009 Aug 13;18(4):301-304. [doi: [10.1007/s10956-009-9172-4](https://doi.org/10.1007/s10956-009-9172-4)]
30. Götz T, Zirngibl A, Pekrun R. Lern- und Leistungsemotionen von Schülerinnen und Schülern. In: Hascher T, editor. *Schule positiv erleben. Erkenntnisse und Ergebnisse zum Wohlbefinden von Schülerinnen und Schülern*. Switzerland: Haupt AG; 2004:49-66.
31. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. *J Med Internet Res* 2011 Dec;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

AR: Augmented Reality

mAR: mobile Augmented Reality

mARble: mobile Augmented Reality blended learning environment

PLRI: Peter L Reichertz Institute for Medical Informatics

POMS: Profile of Mood States questionnaire

SC: single choice

Edited by G Eysenbach; submitted 15.12.12; peer-reviewed by C Brack, S Alipour, C Pimmer; comments to author 22.01.13; revised version received 12.06.13; accepted 29.06.13; published 20.08.13.

Please cite as:

Albrecht UV, Folta-Schoofs K, Behrends M, von Jan U

Effects of Mobile Augmented Reality Learning Compared to Textbook Learning on Medical Students: Randomized Controlled Pilot Study

J Med Internet Res 2013;15(8):e182

URL: <http://www.jmir.org/2013/8/e182/>

doi: [10.2196/jmir.2497](https://doi.org/10.2196/jmir.2497)

PMID: [23963306](https://pubmed.ncbi.nlm.nih.gov/23963306/)

©Urs-Vito Albrecht, Kristian Folta-Schoofs, Marianne Behrends, Ute von Jan. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 20.08.2013. 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.

Viewpoint

The Emergent Discipline of Health Web Science

Joanne S Luciano^{1,2*}, BS, MS, PhD; Grant P Cumming^{3*}, BSc (Hons), MD, FRCOG; Mark D Wilkinson^{4*}, BSc (Hons), PhD; Eva Kahana^{5*}, BA, MA, PhD

¹Web Science Research Center, Tetherless World Constellation, Rensselaer Polytechnic Institute, Troy, NY, United States

²Predictive Medicine, Inc., Belmont, MA, United States

³NHS Grampian, University of Highlands and Islands and University of Aberdeen, Elgin, United Kingdom

⁴Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Madrid, Spain

⁵Elderly Care Research Center, Department of Sociology, Case Western Reserve University, Cleveland, OH, United States

* all authors contributed equally

Corresponding Author:

Joanne S Luciano, BS, MS, PhD

Web Science Research Center

Tetherless World Constellation

Rensselaer Polytechnic Institute

Winslow Building

110 8th St

Troy, NY, 12180

United States

Phone: 1 518 276 4939

Fax: 1 518 276 4464

Email: jluciano@rpi.edu

Abstract

The transformative power of the Internet on all aspects of daily life, including health care, has been widely recognized both in the scientific literature and in public discourse. Viewed through the various lenses of diverse academic disciplines, these transformations reveal opportunities realized, the promise of future advances, and even potential problems created by the penetration of the World Wide Web for both individuals and for society at large. Discussions about the clinical and health research implications of the widespread adoption of information technologies, including the Internet, have been subsumed under the disciplinary label of Medicine 2.0. More recently, however, multi-disciplinary research has emerged that is focused on the achievement and promise of the Web itself, as it relates to healthcare issues. In this paper, we explore and interrogate the contributions of the burgeoning field of Web Science in relation to health maintenance, health care, and health policy. From this, we introduce Health Web Science as a subdiscipline of Web Science, distinct from but overlapping with Medicine 2.0. This paper builds on the presentations and subsequent interdisciplinary dialogue that developed among Web-oriented investigators present at the 2012 Medicine 2.0 Conference in Boston, Massachusetts.

(*J Med Internet Res* 2013;15(8):e166) doi:[10.2196/jmir.2499](https://doi.org/10.2196/jmir.2499)

KEYWORDS

Health Web Science; Medicine 2.0; Web Science; health care singularity; semantic Web; patient engagement; citizen science; crowd-sourcing

Introduction

We present Health Web Science (HWS) as a subdiscipline of Web Science that, while being interested in the Web's impact on health and well-being, also examines the impact of the Web's health-related uses on the design, structure, and evolution of the Web itself. Understanding and appreciating the overlapping yet divergent disciplinary orientation of HWS compared to related research domains motivates specific research efforts

around better utilization of, innovation on, and communication over and within the Web. With this goal in mind, we first introduce the research discipline of Web Science and then describe HWS and its relation to both Web Science and Medicine 2.0. We note that there has been considerable dialogue and controversy in the literature surrounding the now common acceptance of the terms of Medicine 2.0 and Health 2.0 [1,2]. Ultimately recognition of these overlapping fields of inquiry, even in the absence of clear definitions, has had a positive

impact in generating research and policy-relevant dialogue. This paper calls attention to the additional health-relevant intellectual contributions offered by the nascent subdiscipline of HWS. It presents the possibilities and challenges that arise at the intersection between health and the various disciplines related to design, use, and study of the Web. Rather than attempting to formally demarcate boundaries between HWS and Medicine 2.0/Health 2.0, we highlight what we consider to be some of the most useful examples of approaching issues and challenges to improved health and health care as seen through the lens of HWS.

What Is Web Science?

Web Science is the study of the Web as both a social and technical phenomenon: the analysis and synthesis of the World Wide Web and other Web-like information structures, as well as the motivations behind, and social consequences of, these information structures. It encompasses engineering, analytical, methodological, governance, social science, policy, and ethical issues. To study the Web requires extensive collaboration across traditional academic disciplines, building skills and expertise in both the technical underpinnings of the Web and the social processes that have shaped its evolution and its subsequent impact on society [3-6].

From an engineering perspective, issues central to Web Science are largely motivated by requirements generated by emerging technologies such as the global emergence of the Semantic Web, Web Services, consumer-driven content, and peer-to-peer networks. Beyond engineering, however, Web Science studies the Web's topology—its graph-like interconnections and how they can be utilized to understand the nature of the Web itself. From a social science perspective, these same connections represent the overt and concealed interpersonal connections among the Web's users.

The Web Science Research Initiative (WSRI) established in 2006, recognized the importance of the Web and the concept of Web Science. The Web Science Trust (WST) was subsequently founded as a charitable body with the aim of supporting the global development of Web Science. The WST is now a growing international network of world-class research laboratories that supports Web Science research and education, known as the Web Science Network of Laboratories (WSTNet).

What Is Health Web Science?

Health Web Science seeks to understand the interplay between health, health sciences, and the Web, through the academic lens of Web Science. It emerged in response to the juxtaposition of a global health care crisis and an emerging science of the Web. Its roots include a foundational workshop in the summer of 2010 where approximately 40 scholars from different disciplines came together to discuss the influence of the Web on health and how the use of the Web as a health resource has had an impact on the Web itself [7]. Among the participants, there were representatives from diverse practices and scientific communities, including physicians, sociologists, computer scientists, and interested citizens. Participants were in agreement

that the Web has affected, and been affected by, all aspects of health research and delivery and that there is a clear need to unite interests germane to health-related uses of the Web under the rubric of HWS. HWS is, therefore, properly defined as a domain-specific subgroup within Web Science that seeks to understand and describe how the Web shapes, and is shaped by, medicine and health care ecosystems. Through this information, HWS will help engineer the Web and Web-related technologies to facilitate health-related endeavors and empower health professionals, patients, health researchers, and lay communities. Activities relevant to HWS include the synthesis, curation, and discovery of Web pages containing health information; the structure and utilization of interactive social media sites relevant to patient support groups; and semantic annotation and linking of health records and data to facilitate mechanized exploration and analysis. The focus of HWS is, therefore, more broadly aligned with those outside of the medical community and is allied with nonmedical stakeholders disciplines, compared with the interests of Health 2.0 / Medicine 2.0, whose scope reflects more the agendas of professional and particularly medical stakeholders and patients [8].

The motivations to establish Health Web Science as a discipline are multifold. Increasingly, scientific breakthroughs are powered by advanced computing capabilities using the Web as a conduit; therefore, it is important to understand and describe the distinct manner in which the Web is used and engineered for health research, clinical research, and clinical practice. In addition, it is desirable to support consumers who utilize the Web for gathering information about health and well-being and to elucidate approaches to providing social support to both patients and caregivers. Finally, there is the motivation to improve both the effectiveness and efficiency of health care. This is particularly timely since quality improvement and cost containment have become international priorities as governments, employers, and consumers struggle to keep up with rising health care costs. Health Web Science proposes that these motivations will be advanced through a comprehensive study in order to understand the current boundaries and properties of the health-related Web, as well as to inform the design of novel ways to utilize and engineer the Web to maximize its function as a health resource.

In terms of research contributions, HWS-based studies may focus on developing and/or evaluating user responses to Web-based applications that seek to promote "healthy" behaviors, the formation of Web-based communities, or the enabling of Web-based data for consumers to explore and visualize [9]. HWS also promotes research aimed at understanding how humans, individually and in groups, co-create and engage with the Web in areas of their life related to health, medicine, and well-being. Studies anchored in HWS include examination and understanding of "citizen science", discovery of new questions and answers in the Web's metadata, and representation and use of health knowledge by patients, medical professionals, and their machines. HWS researchers would also explore the network effects, the subsequent use/users of the Web, and how policy might need to change given all of these influences [10].

Optimal health models are highly elusive [11], particularly given national and even regional differences in health service needs of diverse populations and funding priorities of different governments and health care organizations. The distributed, adaptable, and highly flexible nature of the Web facilitates the shift from the current model of a centralized, hospital-focused and provider-centric infrastructure, to one where the hospital plays a coordinating role and interacts with the “long tail” of the patient population in a more distributed manner, such as through a peer-to-peer model [12]. Moreover, the Web can play a useful role in tailoring health care to individual needs based not only on medical conditions but also on personal, family, and social factors. Thus, HWS is integral to exploring options and finding solutions to the health problems of the 21st century in both the developing and developed worlds. HWS will enable this shift to a more patient-centric model, as it helps provide the evidence base of which technologies designs and structures work best where and when, under what conditions, and for whom.

We now select examples of HWS research undertakings from several key domains within the discipline, in order to more fully describe the scope, purpose, and impact of HWS.

Web Observatories

The Web Science Trust introduced the concept of a Web observatory as an integrated collection of data sources and data analysis tools that enables observation and experimentation for Web study [13]. The WST’s goal is to mobilize a research community that leverages the strengths of multiple disciplines, methodologies, and theoretical frameworks that research the Web’s infrastructure, transdisciplinary data, visualization, and social networks. This domain is unique to Web Science.

New tools and processes that address the Web’s complexity and multifaceted nature are necessary to build and monitor these Web observatories [9]. Use of semantically enriched tools will facilitate exploration within Web observatories [9]. For example, the new types of data that continue to emerge from both the health and life science domains have enormous potential to improve health and save lives.

Yet, as the US Department of Health and Human Services (HHS) noted in 2011 [14], these data are often inaccessible and incongruent, making them difficult to use. To address these challenges, HHS sought help from the developer community by initiating programming challenges [15,16]. For example, one response that aimed to help improve data accessibility for humans and for machines, enabled the discovery of, access to, and integration of the public health datasets. Making data more accessible empowers individuals to engage. Through the use of tools to aggregate and integrate health information across multiple versions of multiple datasets from multiple source organizations a streamlined, replicable process was created to convert and enhance metadata of the HHS datasets [16-18]. This means that now it becomes easier to create software applications that use the transformed collection of datasets. In one such application, these transformed data were used to help patients decide which hospital was best for them based on personal criteria, such as their medical conditions and

prioritization of factors such as nurse communication and rate of recovery [9].

Social Networks in Health Web Science

Health benefits from social networks. This is not solely a matter of individual responsibility but is influenced by one’s environment, interactions, and by health policy [19]. Health Web Science explores the ways that health-related social networks develop, how they influence health care outcomes, and how these outcomes might be predicted based on the examination of those networks. Ultimately it would be possible to engineer a solution that incorporates the learning from health-related social networks to inform health policy. For example, one might seek to understand the extent to which obesity can be considered contagious [20]. An interested HWS researcher may approach this problem by leveraging a social network analysis model [21] to determine individual network node(s) that are more effective at influencing the whole, or *social cascades*, in which information travels widely through a social network “one hop at a time” through word-of-mouth exchanges between friends. Understanding how these social and online phenomena translate into effects on the health of individual network participants could improve effectiveness through targeted dissemination of health-related messages and crafting of health policy.

Social websites that encourage *participatory cultures* allow communities of patients to enter data on their own medical conditions, providing patients and researchers economies of scale and yielding large datasets that can be mined for patterns that may offer new therapies or change current practices [22]. Such participatory cultures result from the fact that the Web makes it much easier to find those with similar conditions and consequently be able to share insights and search for interventions that appear to be correlated with improved outcomes. Moreover, these self-assembling communities may also act more effectively as lobby groups to demand better products and services to manage their condition or even to influence the direction of health research. Such potential is exemplified in efforts by citizens in Canada to conduct clinical trials for “liberation therapy”, a venoplasty procedure proposed by Zamboni, a vascular surgeon in Italy. Zamboni’s hypothesis is contrary to the accepted view that multiple sclerosis is an autoimmune disorder [23]. His hypothesis has received little attention, except in Canada, where citizens employing social media have demanded clinical trials. This has resulted in heated debates over the Internet, in government, in the media, and in general public discourse, since many health professionals (and citizens) believe there to be no scientific evidence favoring this intervention [23]. It is within the purview of Health Web Science to explore both the positive and negative potential of these self-forming groups, as well as explore policies, protocols, technologies, or infrastructures that may help such groups identify individually appropriate health practices and reject anomalous treatment suggestions.

Patient Engagement Through Citizen Science and Crowdsourcing

Citizen science and crowdsourcing are a form of research collaboration involving members of the public in scientific research to address real-world problems [24]. In health, this type of patient engagement is contrary to the scientific epistemology of evidence-based medicine (EBM), which has been regarded as glacial in its rate of obtaining results [25]. EBM is the conscientious, judicious, and explicit use of current best evidence in making decisions about the care of individual patients [26]. Greenhalgh and Donald [27] tighten the definition of EBM, describing EBM as the use of mathematical estimates of the chance of benefit and the risk of harm derived from high-quality research on population samples to inform clinical decision making. A grassroots alternative strategy to EBM in the collection of population samples, the PatientsLikeMe website exemplifies one Web-based genre of “citizen science”. PatientsLikeMe accumulates and distributes medical data and shares the “research” conducted by its participants.

The citizen science form of evidence gathering has been critiqued by detractors as “full of noise” and “poisonous”, yet, this response to grassroots approaches is akin to the initial resistance toward, but now widespread use of microarray assays [28]. Notably, distributed, human-driven, Web-speed research has shown promise. For example, searching for an association between Gaucher’s disease and Parkinson’s disease using traditional research methodologies took 6 years, whereas using the Web required only 8 months and provided similar results [29]. David deBronkart, diagnosed with kidney cancer and a median survival time at diagnosis of 24 weeks, sought online resources for urgent treatment information. deBronkart obtained his life-saving treatment information in time to save his life, and in gratitude and in recognition of the utility of the Web, deBronkart has become an international spokesperson for patient engagement [30]. Another example of citizen science is exemplified in challenges to lithium’s efficacy in the treatment of amyotrophic lateral sclerosis [31], where citizens experimenting on themselves and reporting their own health data back to the PatientsLikeMe community, discovered that lithium had no effect whatsoever on the disorder—a finding that was subsequently backed up by “traditional” science.

Citizen science, and the latter example in particular, draws attention to issues such as ethics and transparency around data collection and verification methods, as a result of the requirement of individuals to make informed decisions about when and whether to use Web-based information. On the other hand, advocates of obtaining medical advice from peer patient groups say they often learn more than what they get from their doctors because the doctors’ time to read the latest studies, let alone discuss them, is limited [31]; therefore, with respect to informed consent, these patients might consider themselves more informed than traditional clinical-study participants.

Health Web Scientists therefore apply their specific domain-expertise to issues of data and knowledge annotation on the Web, and the technologies and interfaces that facilitate the aggregation, representation, and use of metadata in

supporting patients self-driven health investigations, as well as the ethics and policy around these activities online [32].

Sensors, “Smart” Technologies, and “Expert Patients”

The “quantified self” movement, or “know thyself through numbers” [33], whereby a person collects his/her own personal data from sensors or smartphones and uses this information as a feedback loop for self-improvement, is gathering momentum. This practice aligns with preventative medicine and with the World Health Organization’s (WHO) desire to empower patients to look after and take responsibility for their own health [34]. Many of these applications connect to social networks for, among other things, additional motivational support, eg, websites like RunKeeper or Fitbit. Thus, the quantified self is a natural progression from the current practice of the patient being monitored by health professionals to an individual monitoring themselves.

On a macro level, Health Web Science investigates issues around smart/intelligent cities and, even more broadly, the increasing use of sensors—both fixed and mobile—as ways of monitoring individual and/or community health [35]. All of these use information technologies to create networked infrastructures spanning economies, environments, people, living situations, and governance structures. Similarly, “smart homes” [36] contain Web-based technologies that enable and assist independent living. These technologies have the potential to dramatically increase the amount of data available for analysis and interpretation if it were represented and annotated carefully and made visible over the Web.

As individuals, organizations, populations, and the infrastructure that supports them, become increasingly connected to the Web, Health Web Science has a role in the wellness agenda in a variety of ways. For example, technology can enable patients to remain at home and avoid the expense of hospital-based or nursing home-based care. Networks of “sensors”, both mechanized and human, can provide population-wide surveillance for disease. All of these Web technologies will require deep understanding of the networks they are intended to represent and the health-related requirements placed on these networks by patients, practitioners, and the health care system in general.

In addition to being connected to the Web, through sensors or through networks, individual patients are also connecting to the Web in traditional ways to investigate and evaluate their own health state, particularly if they are suffering from some form of chronic disease. Thus, the concept of the “expert patient” seems closely related to the patient engagement / citizen science topic described above (using the Web socially and as an information source). Here, through a process of research, self-evaluation, and sometimes self-experimentation, guided by both individual exploration as well as social exploration, patients can become highly informed about their own disease. These “expert patients”, through their ability to operate both inside and outside of the health care system, are challenging the traditional relationship between physician and patient [37].

Moreover, these expert patients will soon be sequencing themselves, thereby taking control of even their genomic information out of the hands of traditional care-providers. HWS is already preparing to assist such patients by automatically synthesizing “personalized knowledge-mining workflows”. The Web interface developed by [38] utilizes a combination of semantic text-mining algorithms that extract the concepts from any Web page a patient is reading together with a local database of their personal medical information, eg, drugs prescribed and/or (eventually) personalized DNA sequence. The tool then combines these concepts and data with globally distributed expert knowledge and “mashes up” all of this information to provide personalized hints and guidelines within that Web page. Importantly, health experts, eg, that patient’s clinician, are able to inject their expert opinion into this self-directed exploration by publishing the expert-knowledge models used by the system. Thus, one could imagine that in a future HWS world, the practitioner would provide a computationally readable “health plan” along with a patient’s prescription. The “plan” would explain to the patient’s computer what the treatment trajectory ought to be, and thereby the patient’s Web browser would be able to contextualize the health information content of the Web, based on both the patient’s personal exploration, as well as their clinician’s plan for them. Such focus on patient empowerment through technology provides a striking contrast with the system-centric model of current health care and is strikingly different from the recent pursuit of personalized medicine technologies coming out of genomics and other high-throughput investigations [39,40]. Based upon approaches of Health Web Science, it is reasonable to propose that personalized medicine is going to come from the patient, not the practitioner, reflecting an ethos of patient empowerment facilitated by the Web.

“Big Data”, Semantics, and Other Integration Technologies

The term “big data” refers to datasets that are too large for traditional data management and analytical approaches. Big-data problems emerged as a result of technological advancements that have been rapidly driving the increased size of datasets, even from small or modestly funded laboratories. Thus, the topic of big data is becoming increasingly important to the broader community. HWS proposes that, like the use of metadata in Web sites to support citizen science, metadata supporting the investigation of and integration of big data will enable the discovery of novel treatments for disease, define whole new disease processes, or provide data on drug safety. At the moment, the focus on big data marks a key differentiation between the discipline of Health Web Science and disciplines under the Health 2.0/Medicine 2.0 umbrella. Much work remains to be done in the area of data accessibility, both with data in information silos, as well as data captured using cloud technology, which requires data to be linked and then interrogated in order to be of maximum value. The quantity and complexity of big data necessitates the development of new tools capable of automatically converting this data into new biomedical knowledge that is accessible to the clinical practitioners, researchers, and the public, over the Web. It is in this context that the emergence of the Semantic Web, and its

related technologies, are important. Unlike the traditional Web, where connections between data elements were based on human-readable hyperlinks, on the Semantic Web, these linkages become machine-readable. Thus, machines can explore vast networks of interconnected data in a meaningful and computationally efficient way.

These new opportunities also pose new challenges. HWS practitioners must address issues related to data sharing and discovery, data interoperability, knowledge representation, and exchange. Privacy issues must also be addressed using skills of HWS practitioners. Web architecture can ultimately enable better-connected health information, patients, practitioners, and health researchers [41].

Rapid, Automated, Contextualized Knowledge Discovery and Application

Clinicians and health researchers are struggling to keep up to date with the medical literature and best practices despite specialization and subspecialization of journals. Today, a typical primary care doctor must stay abreast of approximately 10,000 diseases and syndromes, 3000 medications, and 1100 laboratory tests, and the list grows every year [42]. This information must somehow be filtered and presented to the right clinician at the right time. Moreover, the knowledge embedded in this voluminous literature must be applied judiciously and specifically. This presents an opportunity for Health Web Technologies. Health Web Science proposes that the careful, automated integration of this voluminous knowledge, in the context of the individualized data of a particular patient, will enable the promise of efficiency and quality to finally be delivered—the right treatment, at the right time, for the right patient.

Through novel Web technologies, the discovery and application of new medical knowledge will become increasingly instantaneous, a point described by some as “the health care singularity” [42]. Accordingly, when new information is generated through social networks, clinical research, or “bench-science”, its subsequent translation will affect treatment at the bedside and has the potential to become instantaneous.

Due to the size and complexity of big data, particularly in the era of personalized medicine (eg, personal genomes), clinical decision making will need to be increasingly mechanized. Humans, no matter how well trained, cannot usefully process datasets of this scale. As such, Semantic Web technologies will increasingly play a role in ensuring that this new knowledge and its influence on clinical decision making of the machine can be immediately understood by practitioners through the accurate communication of the rationale for machine-recommended decisions.

Far from replacing the expert, this ensures that the clinician maintains their critical final-arbiter role in patient care. This is of utmost importance and is one of the primary reasons that health professionals undergo a long period of training. Aristotle referred to it as “phronesis” or practical wisdom [43]. Human beings reason differently from computers and possess the ability to take into account factors that computers cannot currently

perceive [44] because the logic of care is unbounded, nonlinear, and unpredictable. There cannot be an algorithm for every situation. This debate of “intuition versus formula” [45] and “expert intuition, can we trust it” [46] will continue as semantic and artificial intelligence technologies expand and mature. As key users of, and proponents of, these technologies, Health Web Science practitioners need to be involved in this debate providing insight and balance.

From a technology perspective, the outbreak of severe acute respiratory syndrome (SARS) in 2002–2003 provides an example of Health Web Science in action. SARS was characterized by 3 phenomena: how fast the disease spread, how quickly it was stopped, and how a virulent virus killed relatively few people. The WHO used the Global Public Health Intelligence Network software tool developed by Canada’s National Health Ministry to scan newswires and Internet sources for mentions of possible (SARS) outbreaks or other unexplained health events. More than one third of the rumors identified by the tool led the WHO to identify and isolate cases of SARS [47].

Conclusion

We present Health Web Science as subdiscipline—a subdiscipline of the new field of Web Science that focuses on the mutual interplay between the World Wide Web, the health data it must contain, and all who utilize it. Health Web Science complements and overlaps with the discipline of Medicine 2.0, but differs in a focus on the role of the Web in health outcomes. Health care delivery is undergoing a revolutionary shift as knowledge is decentralized. The doctor-patient relationship is

one illustration, that is, from a doctor-knows-best top-down informed model to a shared decision-making model. This transformation is achieved through the Web of linked documents and patient utilization of the Social Web. The results can be easily witnessed in the growth of “patient power” and the increased influence of patient groups in exchanges with health professionals. Such advances facilitate sharing of both positive and negative patient experiences and serve to disseminate information much more efficiently. The speed of change in technological advances is exponential, resulting in the predicted achievement of the “health care singularity” in which information flow from research to practice is instantaneous.

HWS, therefore, has a role to play in explicating the Web aspects contributing to a personalized, predictive, preventative, and participatory medicine. These contributions occur in the context of the technological intersection between medical experts, expert patients, and increasingly rapid knowledge dissemination. It has the potential to unlock the secrets of big data within a carefully controlled governance framework and help in separating Web fact from Web fiction. It can transform the generalized nature of Web information, making it relevant and applicable to an individual patient. HWS is the pursuit of novel ways to provide relevant, accurate, personalized, expert medical information, and evidence-based guidance to patients as they manage their own health care. The momentum behind this emergent discipline has been established, and dialogue now needs to continue so that the various stakeholder communities can mutually educate each other for the greater benefit of society.

Acknowledgments

This work was partially supported by the National Health Service Grampian, the University of the Highlands and Islands, the Center for Plant Biotechnology and Genomics at UPM/INIA, Case Western Reserve University, and the Tetherless World Constellation at Rensselaer Polytechnic Institute. The ideas presented at Medicine 2.0 and in this paper have benefited greatly from discussions with following people: Christopher JO Baker, Deborah McGuiness, David Molik, Catherine Pope, Tracy Ann Kosa, John Wilbanks, Doug McKendrick, Elizabeth H Brooks, Marie Joan Kristine Gloria, and Holly Jarmin.

Conflicts of Interest

None declared.

References

1. 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](https://pubmed.ncbi.nlm.nih.gov/18725354/)]
2. Hughes B, Joshi I, Wareham J. Health 2.0 and Medicine 2.0: tensions and controversies in the field. *J Med Internet Res* 2008 Aug;10(3):e23 [FREE Full text] [doi: [10.2196/jmir.1056](https://doi.org/10.2196/jmir.1056)] [Medline: [18682374](https://pubmed.ncbi.nlm.nih.gov/18682374/)]
3. Web Science. University of Southampton, UK URL: <http://webscience.ecs.soton.ac.uk/> [accessed 2012-12-15] [WebCite Cache ID 6CwGSQatt]
4. Berners-Lee T, Weitzner DJ, Hall W, O'Hara K, Shadbolt N, Hendler JA. A Framework for Web Science. *FNT in Web Science* 2006;1(1):1-130. [doi: [10.1561/1800000001](https://doi.org/10.1561/1800000001)]
5. Berners-Lee T, Hall W, Hendler J, Wetzner D. Creating a Science of the Web. *Science* 2006 Aug 11;313(5788):769-771. [doi: [10.1126/science.1126902](https://doi.org/10.1126/science.1126902)]
6. Fensel D. *Spinning the semantic web: bringing the World Wide Web to its full potential*. Cambridge, Mass: MIT Press; 2003.
7. Brooks EH, Cumming GP, Luciano JS. Health Web science: application of Web science to the area of health education and health care. In: *Proceedings of the Second international workshop on Web science and information exchange in the*

- medical Web. 2011 Presented at: Second international workshop on Web science and information exchange in the medical web; October 28, 2011; Glasgow, UK p. 11-14.
8. Van De Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. *J Med Internet Res* 2010;12(2):e18 [FREE Full text] [doi: [10.2196/jmir.1350](https://doi.org/10.2196/jmir.1350)] [Medline: [20542857](https://pubmed.ncbi.nlm.nih.gov/20542857/)]
 9. Gloria MJK, Luciano JS, McGuinness DL. Building Semantically-Enriched Web Observatories. In: Proceedings of Web Science Conference. 2013 Presented at: Web Science Conference; May 2-4, 2013; Paris, France.
 10. Pariser E. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press HC; 2011.
 11. McGlynn EA. There is no perfect health system. *Health Aff (Millwood)* 2004;23(3):100-102 [FREE Full text] [Medline: [15160807](https://pubmed.ncbi.nlm.nih.gov/15160807/)]
 12. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 15;328(7449):1166 [FREE Full text] [doi: [10.1136/bmj.328.7449.1166](https://doi.org/10.1136/bmj.328.7449.1166)] [Medline: [15142921](https://pubmed.ncbi.nlm.nih.gov/15142921/)]
 13. Hall W, Tiropanis T. Web evolution and Web science. *Computer Networks* 2012;56(18):3859-3865.
 14. U.S. Health and Human Services and The Office of the National Coordinator for Health Information Technology introduce new Investing in Innovations (i2) Initiative. Washington, DC: HHS Press Office; 2011 Jun 8. URL: <http://www.hhs.gov/news/press/2011pres/06/20110608a.html> [accessed 2013-06-05] [WebCite Cache ID 6H9GDhcu5]
 15. US Department of Health and Human Services, Office of the National Coordinator for Health IT. Health Data Platform Metadata Challenge. 2013. URL: <http://www.health2con.com/devchallenge/health-data-platform-metadata-challenge/> [accessed 2013-04-07] [WebCite Cache ID 6Fhcl51GI]
 16. Wong A. Health 2.0 website. 2013 Feb 20. Winners of Health Data Platform Challenge URL: <http://www.health2con.com/news/2013/02/20/winners-of-health-data-platform-challenges/> [accessed 2013-06-05] [WebCite Cache ID 6H9GNdrj5]
 17. McCusker J, Lebo T, Graves A. Health Data Statistics. 2013. URL: <http://healthdata.tw.rpi.edu/statistics> [accessed 2013-06-05] [WebCite Cache ID 6H9GRj7RC]
 18. McCusker J, Lebo T, Graves A, Gloria M. Wiki: TWC-HealthData. 2013. URL: <https://github.com/jimmccusker/twc-healthdata/wiki> [accessed 2013-06-05] [WebCite Cache ID 6H9GaCQ8x]
 19. The PLoS Medicine Editors. Social Relationships Are Key to Health, and to Health Policy. *PLoS Med* 2010;7(8):e1000334. [doi: [10.1371/journal.pmed.1000334](https://doi.org/10.1371/journal.pmed.1000334)]
 20. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007 Jul 26;357(4):370-379. [doi: [10.1056/NEJMsa066082](https://doi.org/10.1056/NEJMsa066082)] [Medline: [17652652](https://pubmed.ncbi.nlm.nih.gov/17652652/)]
 21. Cha M, Mislove A, Gummadi KP. A measurement-driven analysis of information propagation in the flickr social network. In: Proceedings of the 18th international conference on World Wide Web. 2009 Presented at: 18th International Conference on World Wide Web; April 20-24, 2009; Madrid, Spain p. 721-730.
 22. Chafe R, Born KB, Slutsky AS, Laupacis A. The rise of people power. *Nature* 2011;472(7344):410. [doi: [10.1038/472410a](https://doi.org/10.1038/472410a)]
 23. Moran L. Canadian Government Allocates Funds to Investigate "Liberation Treatment" for Multiple Sclerosis: Sacrificing Science at the Altar of Consumer Demand. 2012 Jun 07. URL: <http://sandwalk.blogspot.com.es/2012/06/canadian-government-allocates-funds-to.html> [accessed 2012-12-15] [WebCite Cache ID 6CwGV0gzY]
 24. Wiggins A, Crowston K. From conservation to crowdsourcing: A typology of citizen science. In: Proceedings of the Hawaii International Conference on System Sciences. 2011 Presented at: 44th Hawaii International Conference on System Sciences; Jan. 4-7, 2011; Kauai p. 1-10.
 25. Goetz T. *Wired Magazine*. 2010. Sergey Brin's Search for a Parkinson's Cure URL: http://www.wired.com/magazine/2010/06/ff_sergeys_search/ [accessed 2013-06-05] [WebCite Cache ID 6H9HbJdzg]
 26. Sackett D, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996 Jan 13;312(7023):71-72 [FREE Full text] [Medline: [8555924](https://pubmed.ncbi.nlm.nih.gov/8555924/)]
 27. Greenhalgh T. *How to Read a Paper: The Basics of Evidence-Based Medicine (HOW - How To)*. Hoboken, NJ: Wiley-Blackwell - BMJ Books; 2006.
 28. Editorial. Microarray policy. *Nat Immunol* 2003;4(93):---. [doi: [10.1038/ni0203-93](https://doi.org/10.1038/ni0203-93)]
 29. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol* 2011 May;29(5):411-414. [doi: [10.1038/nbt.1837](https://doi.org/10.1038/nbt.1837)] [Medline: [21516084](https://pubmed.ncbi.nlm.nih.gov/21516084/)]
 30. deBronkart D, Kuraitis V, Kibbe D. Gimme my damn data!. URL: <http://thehealthcareblog.com/blog/2010/05/07/gimme-my-damn-data/> [accessed 2013-06-05] [WebCite Cache ID 6H9IFYBBO]
 31. Shute N. Getting Medical Advice on the Web from Other Patients: Patients swap personal stories, offer help, and share studies at "crowd-sourced" disease sites. 2009 Jul 13. URL: <http://health.usnews.com/health-news/best-hospitals/articles/2009/07/13/getting-medical-advice-on-the-web-from-other-patients> [accessed 2013-06-05] [WebCite Cache ID 6H9ITqqLe]
 32. Taylor D, Bury M. Chronic illness, expert patients and care transition. *Sociol Health Illn* 2007 Jan;29(1):27-45. [doi: [10.1111/j.1467-9566.2007.00516.x](https://doi.org/10.1111/j.1467-9566.2007.00516.x)] [Medline: [17286704](https://pubmed.ncbi.nlm.nih.gov/17286704/)]
 33. Wolf GI. AETHER, Quantified Self. URL: <http://aether.com/quantifiedself> [accessed 2012-03-27] [WebCite Cache ID 66TEHdz4d]

34. Commission of the European Communities. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions eEurope 2002: Quality Criteria for Health-related Websites. 2002. URL: http://www.hon.ch/HONcode/HON_CCE_en.htm [accessed 2013-06-05] [WebCite Cache ID 6H9InXmDT]
35. WHO/Europe. Empowering patients.: World Health Organization Regional Office for Europe URL: <http://www.euro.who.int/en/what-we-do/health-topics/noncommunicable-diseases/sections/news/2012/4/empowering-patients> [accessed 2013-06-05] [WebCite Cache ID 6H9It5u99]
36. Smart Cities Project. Investing in the future by working together for a sustainable and competitive region. The North Sea Region Programme URL: <http://www.smartcities.info/> [accessed 2012-12-15] [WebCite Cache ID 6CwGXQBK5]
37. Morris M, Ozanne E, Miller K, Santamaria N, Pearce A, Said C, et al. Smart technologies for older people: a systematic literature review of smart technologies that promote health and wellbeing of older people living at home.: Institute for a Broadband-Enabled Society; 2012. URL: <http://broadband.unimelb.edu.au/health/ageing/smart-technologies-for-older-people.html> [WebCite Cache ID 6CwGZwwz2]
38. Vandervalk B, McCarthy EL, Cruz-Toledo J, Klein A, Baker CJ, Dumontier M, et al. The SADI Personal Health Lens: A Web Browser-Based System for Identifying Personally Relevant Drug Interactions. JMIR Res Protoc 2013;2(1):e14. [doi: 10.2196/resprot.2315] [Medline: 23612187]
39. S. 3822 - 109th Congress: Genomics and Personalized Medicine Act of 2006. 2006. URL: <http://www.govtrack.us/congress/bills/109/s3822> [accessed 2013-06-05] [WebCite Cache ID 6H9JNFyvy]
40. S. 976 - 110th Congress: Genomics and Personalized Medicine Act of 2007. 2007. URL: <http://www.govtrack.us/congress/bills/110/s976> [accessed 2013-06-05] [WebCite Cache ID 6H9JPlobn]
41. Ranck J. Connected Health: How Mobile Phones, Cloud and Big Data Will Reinvent Healthcare. online: GigOM Books; 2012.
42. Gillam M, Feied C, Handler J, Moody E, Shneiderman B, Plaisant C, et al. The healthcare singularity and the age of semantic medicine. In: The Fourth Paradigm: Data-Intensive Scientific Discovery. online: Microsoft Research; 2009:57-64.
43. Phronēsis (Greek: φρόνησις) is an Ancient word for wisdom or intelligence which is a common topic of discussion in philosophy. URL: <http://en.wikipedia.org/wiki/Phronesis> [accessed 2012-12-19] [WebCite Cache ID 6D1vpHMxY]
44. Greenhalgh T. Why do we always end up here? Evidence-based medicine's conceptual cul-de-sacs and some off-road alternative routes. J Prim Health Care 2012 Jun;4(2):92-97. [Medline: 22675691]
45. Meehl P. Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence (The Master Work Series). Minnesota: Jason Aronson; 1954.
46. Kahneman D. Thinking, fast and slow. Toronto: Penguin Books Ltd; 2011.
47. Zuckerman E. Rewire: Digital cosmopolitans in the age of connection. New York: W.W. Norton & Company; 2013.

Abbreviations

- DNA:** Deoxyribonucleic acid
EBM: Evidence Based Medicine
HHS: Health and Human Services
HWS: Health Web Science
SARS: severe acute respiratory syndrome
WHO: World Health Organization
WSRI: Web Science Research Institute
WSTNet: Web Science Trust Network of Labs

Edited by G Eysenbach; submitted 16.12.12; peer-reviewed by EK Yun, YH Weng; comments to author 28.01.13; revised version received 09.04.13; accepted 24.04.13; published 22.08.13.

Please cite as:

Luciano JS, Cumming GP, Wilkinson MD, Kahana E
The Emergent Discipline of Health Web Science
J Med Internet Res 2013;15(8):e166
URL: <http://www.jmir.org/2013/8/e166/>
doi:10.2196/jmir.2499
PMID:23968998

©Joanne S Luciano, Grant P Cumming, Mark D Wilkinson, Eva Kahana. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 22.08.2013. 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.

Letter to the Editor

Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research

John A Cunningham^{1,2}

¹Centre for Mental Health Research, The Australian National University, Canberra, Australia

²Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:

John A Cunningham

Centre for Mental Health Research

The Australian National University

Building 63,

Canberra, 0200

Australia

Phone: 61 6125 1859

Fax: 61 6125 0733

Email: john.cunningham@anu.edu.au

(*J Med Internet Res* 2013;15(8):e179) doi:[10.2196/jmir.2845](https://doi.org/10.2196/jmir.2845)

KEYWORDS

research methods; randomised controlled trials; alcohol; Internet; brief intervention

One of the challenging issues facing the randomized controlled trial (RCT) researcher is how to interpret the results of studies where there are improvements in the behaviour under study but where the degree of improvement does not differ between the experimental conditions [1]. This is especially a challenge when the RCT involves the comparison of two or more interventions rather than an intervention compared to some form of no-intervention control group.

One possible cause of the observed improvement in such trials is that both interventions were “active” - that both interventions were effective in facilitating or causing a change among participants. Unfortunately, there is no way to determine if this claim is definitely true from the results of a negative RCT. Other interpretations of the results include: 1) that the change over time is due to regression to the mean [2, 3]; 2) due to natural history maturation (meaning that participants were in a period in their lives where, on average, a downward trend in quantity of drinking could be expected); or 3) the trial recruited participants who were already motivated to change and who

would have done so anyway without exposure to the interventions under study [4].

Any of these alternate explanations could apply to the recent trial by Hester and colleagues [1]. Further, there is a well-established finding in the alcohol research field that participants in the no intervention control condition of intervention trials show improvements in their drinking from baseline to follow-up [5]. This may be particularly the case in trials recruiting participants from the community rather than from treatment settings where intractable alcohol problems are more common [6]. Essentially, the assumption that any changes over time are due to the intervention in a negative trial is predicated on the assumption that the participants would show no improvement without receiving some type of intervention. There may be some behaviours where this is the case. However, alcohol abuse is demonstrably not one of them. Thus, it is unwise to favour an intervention effect explanation over other causes when faced with the results of an RCT where participants show improvement over time but that there are no significant statistical differences between intervention conditions.

References

1. Hester RK, Lenberg KL, Campbell W, Delaney HD. Overcoming Addictions, a Web-Based Application, and SMART Recovery, an Online and In-Person Mutual Help Group for Problem Drinkers, Part 1: Three-Month Outcomes of a Randomized Controlled Trial. *J Med Internet Res* 2013;15(7):e134 [FREE Full text] [doi: [10.2196/jmir.2565](https://doi.org/10.2196/jmir.2565)] [Medline: [23846588](https://pubmed.ncbi.nlm.nih.gov/23846588/)]
2. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol* 2005 Feb;34(1):215-220 [FREE Full text] [doi: [10.1093/ije/dyh299](https://doi.org/10.1093/ije/dyh299)] [Medline: [15333621](https://pubmed.ncbi.nlm.nih.gov/15333621/)]

3. Morton V, Torgerson DJ. Effect of regression to the mean on decision making in health care. *BMJ* 2003 May 17;326(7398):1083-1084 [[FREE Full text](#)] [doi: [10.1136/bmj.326.7398.1083](https://doi.org/10.1136/bmj.326.7398.1083)] [Medline: [12750214](https://pubmed.ncbi.nlm.nih.gov/12750214/)]
4. Cunningham JA. Regression to the mean: what does it mean? *Alcohol Alcohol* 2006 Oct;41(5):580; author reply 581 [[FREE Full text](#)] [doi: [10.1093/alcalc/agl039](https://doi.org/10.1093/alcalc/agl039)] [Medline: [16751218](https://pubmed.ncbi.nlm.nih.gov/16751218/)]
5. Moyer A, Finney JW. Outcomes for untreated individuals involved in randomized trials of alcohol treatment. *J Subst Abuse Treat* 2002 Oct;23(3):247-252. [Medline: [12392812](https://pubmed.ncbi.nlm.nih.gov/12392812/)]
6. Cunningham JA, McCambridge J. Is alcohol dependence best viewed as a chronic relapsing disorder? *Addiction* 2012 Jan;107(1):6-12 [[FREE Full text](#)] [doi: [10.1111/j.1360-0443.2011.03583.x](https://doi.org/10.1111/j.1360-0443.2011.03583.x)] [Medline: [21981681](https://pubmed.ncbi.nlm.nih.gov/21981681/)]

Edited by G Eysenbach; submitted 23.07.13; peer-reviewed by A Moyer, R Hester; accepted 15.08.13; published 19.08.13.

Please cite as:

Cunningham JA

Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research

J Med Internet Res 2013;15(8):e179

URL: <http://www.jmir.org/2013/8/e179/>

doi: [10.2196/jmir.2845](https://doi.org/10.2196/jmir.2845)

PMID: [23958686](https://pubmed.ncbi.nlm.nih.gov/23958686/)

©John A Cunningham. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 19.08.2013. 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.

Letter to the Editor

Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research - Author's Reply

Reid Hester¹, PhD (Clinical Psycho); William Campbell¹, MS(Psycho); Kathryn Lenberg², PhD (Clin.Psycho.); Harold Delaney³, PhD(Psycho)

¹Behavior Therapy Associates, LLC, Research Div., Albuquerque, NM, United States

²Presbyterian Medical Group, Albuquerque, NM, United States

³University of New Mexico, Dept. of Psychology, Albuquerque, NM, United States

Corresponding Author:

Reid Hester, PhD (Clinical Psycho)
Behavior Therapy Associates, LLC
Research Div.
9426 Indian School Rd NE Ste 1
Albuquerque, NM, 87112
United States
Phone: 1 5053456100
Fax: 1 5053454531
Email: reidkhester@gmail.com

(*J Med Internet Res* 2013;15(8):e180) doi:[10.2196/jmir.2884](https://doi.org/10.2196/jmir.2884)

KEYWORDS

addictions; cognitive-behavioral program; Web application; SMART Recovery; mutual self-help groups

In Response

Cunningham is correct in noting that it is difficult to draw conclusions about the results of a randomized clinical trial (RCT) when two or more active interventions are compared without utilizing a no-treatment control condition. This is an issue that bedevils clinical research, but it is also one that, ethically speaking, has long been resolved in the area of addictions. As noted in the *New England Journal of Medicine*, when evaluating the risk-benefit ratio of employing a placebo or no-treatment control, potential psychological and social harms must be addressed [1]. The participants in this study comprised some of the most severe drinkers in our 20-plus years of conducting RCTs of computer delivered interventions for problem drinkers. Moreover, in this case, we were recruiting individuals in the contemplation and active stages of change, a window of opportunity well recognized among clinicians as the time when an intervention has the best chance of helping. Since treatment-seeking individuals struggling with serious alcohol disorders may be harmed by temporary conditions, such as a wait-list, the ethical criterion of doing no harm would not be met. Thus the clinically relevant question for us clearly became not whether a new treatment is better than nothing, but whether it is better than another treatment.

Given this limitation, it is incumbent upon researchers to be clear about their methods, conservative in their analysis, and

parsimonious in their conclusions. To this end, we were clear when addressing our limitations that the lack of a no-treatment comparison group “prevents us from being assured that the treatment assigned was the cause of the improvement.” Cunningham is correct in observing that, as in virtually all clinical trials involving active interventions, alternative explanations may possibly account for the changes we observed in this trial. He mentions, in particular, client motivation and regression to the mean, so we will address those two concerns here.

While natural recovery does occur for many people with alcohol problems, it typically does so for those who tend to be at the less severe end of the spectrum [2]. As we noted, that would not likely pertain to the sample studied here, whose mean AUDIT scores were 24.7 and InDuC scores were 41.4. It is relevant to note that the mean within-group effect size observed in no-treatment control groups is typically much smaller than the d of .97 seen in the OA+SR group or .96 in the SR only group in our study. For example, the comparable mean pre-post effect sizes in five control groups reported in White et al. [3] ranged from -.61 to +.22 for drinks consumed with an overall mean weighted by study size of only $d = +.05$. Thus, while other explanations of the cause of the effects we observed are conceivable in theory, in practice the implemented treatments seemed the most plausible explanation for the large effects observed.

As for the participants' level of motivation coming in to the study, we were clear in our methods that we recruited from individuals who were actively seeking treatment options, and in so doing had arrived at the SMART Recovery website. This does bespeak of motivation, it's hard to deny; but it is also hard to argue that we would have obtained the results we did in this clinical trial on the basis of motivation alone. It is a clinical truism that behavioral interventions require motivation on the part of patients to be effective. To that end, our results showed that whether participants in the trial received that intervention from SMART Recovery groups or from the Overcoming Addictions app, they generally succeeded in making significant reductions in their drinking and alcohol-related problems. That the data did not support better outcomes in one group or the other means, to us, that problem drinkers looking for help becoming and remaining abstinent, have options which are equally effective. And indeed, this is all we claim in our conclusions.

Cunningham characterizes the outcomes of our RCT as a "negative" trial and concludes that it is "unwise to favour an intervention effect explanation over other causes when faced with the results of an RCT where participants show improvement over time but that there are no significant differences between intervention conditions." While the lack of difference between conditions does in fact remain a topic of empirical interest, we find his use of the term "negative" baffling, given the highly positive changes across both groups. We are also curious why Cunningham failed to raise these concerns earlier, with a plethora of other studies that compared active treatments without a no-treatment control [4, 5, 6, 7, 8] including his own [9].

Although we believe that Cunningham's concerns with our conclusions are overblown, we welcome the opportunity to discuss the research methods of web-based interventions. Cunningham has argued convincingly elsewhere about the particular difficulties of research in this area [10], and to that end we strove as much as possible to generate results that would be generalizable. In this case developing an application specifically for individuals seeking help, and engaging in the intervention, on-line. Also, as per his recommendation [10], we sought to concurrently evaluate the effectiveness of SMART Recovery, the protocol upon which our application was based, since this had not been done previously.

Returning to the question of what factors mediate the effectiveness of two treatments under comparison, it might be interesting for Cunningham to explain how his own study [9], that found added benefits in a more involved intervention for problem drinkers, compared to Hansen et al.'s, findings which showed no such added benefit [8]. It is actually such questions as these that are more pressing for research into web-based interventions than are questions about the implications of not using a no-treatment control condition. As a field, we lack insight into these and other questions, such as why web-based interventions work in some instances and not others, why increased engagement leads to better outcomes in some cases and not others, and what sorts of individuals are most likely to benefit from web-based interventions. These are the sorts of data that other researchers in our field need, and it is just such data we will be reporting in Part 2 of our paper along with our six-month follow-up results.

References

1. Emanuel E, Miller FG. The ethics of placebo-controlled trials--a middle ground. *N Engl J Med* 2001 Sep 20;345(12):915-919. [doi: [10.1056/NEJM200109203451211](https://doi.org/10.1056/NEJM200109203451211)] [Medline: [11565527](https://pubmed.ncbi.nlm.nih.gov/11565527/)]
2. Klingemann H, Sobell LC. Promoting self-change from addictive behaviors practical implications for policy, prevention, and treatment. New York: Springer; 2007.
3. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. *J Med Internet Res* 2010;12(5):e62 [FREE Full text] [doi: [10.2196/jmir.1479](https://doi.org/10.2196/jmir.1479)] [Medline: [21169175](https://pubmed.ncbi.nlm.nih.gov/21169175/)]
4. Leykin Y, Aguilera A, Torres LD, Pérez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an International Smoking Cessation Study. *J Med Internet Res* 2012;14(1):e5 [FREE Full text] [doi: [10.2196/jmir.1829](https://doi.org/10.2196/jmir.1829)] [Medline: [22314016](https://pubmed.ncbi.nlm.nih.gov/22314016/)]
5. 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/)]
6. van Genugten L, van Empelen P, Boon B, Borsboom G, Visscher T, Oenema A. Results from an online computer-tailored weight management intervention for overweight adults: randomized controlled trial. *J Med Internet Res* 2012;14(2):e44 [FREE Full text] [doi: [10.2196/jmir.1901](https://doi.org/10.2196/jmir.1901)] [Medline: [22417813](https://pubmed.ncbi.nlm.nih.gov/22417813/)]
7. Collins CE, Morgan PJ, Jones P, Fletcher K, Martin J, Aguiar EJ, et al. A 12-week commercial web-based weight-loss program for overweight and obese adults: randomized controlled trial comparing basic versus enhanced features. *J Med Internet Res* 2012;14(2):e57 [FREE Full text] [doi: [10.2196/jmir.1980](https://doi.org/10.2196/jmir.1980)] [Medline: [22555246](https://pubmed.ncbi.nlm.nih.gov/22555246/)]
8. Hansen AB, Becker U, Nielsen AS, Grønbaek M, Tolstrup JS, Thygesen LC. Internet-based brief personalized feedback intervention in a non-treatment-seeking population of adult heavy drinkers: a randomized controlled trial. *J Med Internet Res* 2012;14(4):e98 [FREE Full text] [doi: [10.2196/jmir.1883](https://doi.org/10.2196/jmir.1883)] [Medline: [22846542](https://pubmed.ncbi.nlm.nih.gov/22846542/)]
9. Cunningham JA. Comparison of two internet-based interventions for problem drinkers: randomized controlled trial. *J Med Internet Res* 2012;14(4):e107 [FREE Full text] [doi: [10.2196/jmir.2090](https://doi.org/10.2196/jmir.2090)] [Medline: [22954459](https://pubmed.ncbi.nlm.nih.gov/22954459/)]
10. Cunningham JA. Internet evidence-based treatments. In: Miller P, editor. Evidence Based Addiction Treatment. Amsterdam: Elsevier; 2009:379-398.

Edited by G Eysenbach; submitted 15.08.13; this is a non-peer-reviewed article; accepted 15.08.13; published 19.08.13.

Please cite as:

Hester R, Campbell W, Lenberg K, Delaney H

Claiming Positive Results From Negative Trials: A Cause for Concern in Randomized Controlled Trial Research - Author's Reply
J Med Internet Res 2013;15(8):e180

URL: <http://www.jmir.org/2013/8/e180/>

doi: [10.2196/jmir.2884](https://doi.org/10.2196/jmir.2884)

PMID: [23958742](https://pubmed.ncbi.nlm.nih.gov/23958742/)

©Reid Hester, William Campbell, Kathryn Lenberg, Harold Delaney. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 19.08.2013. 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/>